

Notes for *Foundations of Modern Analysis* by
Avner Friedman

Anton Ottosson – antonott@kth.se

January 4, 2025

Chapter 1 – Measure Theory

Section 1.1 – Rings and Algebras

Problems

1.1.1

$$\left(\varliminf_n E_n\right)^c = \overline{\varliminf_n E_n^c}, \quad \left(\overline{\varliminf_n E_n}\right)^c = \varliminf_n E_n^c.$$

Solution. Note that

$$\begin{aligned} x \in \varliminf_n E_n &\iff x \in E_n \text{ for all but finitely many } n \\ &\iff x \in E_n^c \text{ for only finitely many } n. \end{aligned}$$

Hence

$$\begin{aligned} x \in \left(\varliminf_n E_n\right)^c &\iff x \in E_n^c \text{ for infinitely many } n \\ &\iff x \in \overline{\varliminf_n E_n^c}, \end{aligned}$$

proving the first identity.

Next, let $F_n = E_n^c$ for every n . Then

$$\overline{\varliminf_n E_n} = \overline{\varliminf_n F_n^c} = \left(\varliminf_n F_n\right)^c = \left(\varliminf_n E_n^c\right)^c$$

by the first identity, and the second identity follows.

1.1.2

$$\overline{\varliminf_n E_n} = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n, \quad \varliminf_n E_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n.$$

Solution. Suppose $x \in \overline{\lim}_n E_n$. Then $x \in E_n$ for infinitely many n . It follows that $x \in \bigcup_{n=k}^{\infty} E_n$ for all $k \in \mathbb{N}$, and hence that $x \in \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n$.

Conversely, assume that $x \in \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n$. Then $x \in \bigcup_{n=k}^{\infty} E_n$ for all $k \in \mathbb{N}$. It follows that $x \in E_n$ for infinitely many n , and thus that $x \in \overline{\lim}_n E_n$. This proves the first identity.

Next, suppose that $x \in \underline{\lim}_n E_n$. Then $x \in E_n$ for all but finitely many n , so there is some $k' \in \mathbb{N}$ such that $x \in E_n$ for all $n \geq k'$. It follows that $x \in \bigcap_{n=k'}^{\infty} E_n$, and hence that $x \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n$.

Conversely, assume that $x \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n$. Then $x \in \bigcap_{n=k'}^{\infty} E_n$ for some $k' \in \mathbb{N}$, which means that $x \in E_n$ for all $n \geq k'$. It follows that $x \in E_n$ for all but finitely many n ; that is, $x \in \underline{\lim}_n E_n$.

1.1.3

If \mathcal{R} is a σ -ring and $E_n \in \mathcal{R}$, then

$$\bigcap_{n=1}^{\infty} E_n \in \mathcal{R}, \quad \overline{\lim}_n E_n \in \mathcal{R}, \quad \underline{\lim}_n E_n \in \mathcal{R}.$$

Solution. Let $Y = \bigcup_{n=1}^{\infty} E_n \in \mathcal{R}$. Then $E_n \subset Y$ for all n , and it follows that

$$\bigcap_{n=1}^{\infty} E_n = Y \cap \left(\bigcap_{n=1}^{\infty} E_n \right) = Y - \left(Y - \bigcap_{n=1}^{\infty} E_n \right).$$

Notice that

$$Y - \bigcap_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} (Y - E_n) \in \mathcal{R},$$

by properties (b) and (e). (The equality is analogous to the identity (1.1.2), but with Y in place of X .) It follows (by (b) again) that $\bigcap_{n=1}^{\infty} E_n \in \mathcal{R}$. For later reference, let us call this result (x).

Given $k \in \mathbb{N}$, let $A_n = \emptyset$ for $n < k$, and let $A_n = E_n$ for $n \geq k$. Then $A_n \in \mathcal{R}$ for all n by (a), hence

$$\bigcup_{n=k}^{\infty} E_n = \bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$$

by (e). It then follows by (x) that

$$\overline{\lim}_n E_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n \in \mathcal{R}.$$

By a similar argument we find that (x) implies

$$\bigcap_{n=k}^{\infty} E_n \in \mathcal{R}$$

for all $k \in \mathbb{N}$. Thus

$$\varliminf_n E_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n \in \mathcal{R}$$

by (e).

1.1.4

The intersection of any collection of rings (algebras, σ -rings, or σ -algebras) is also a ring (an algebra, σ -ring, or σ -algebra).

Solution. Let \mathcal{C} be a collection of classes. Let $\bigcap \mathcal{C}$ denote the intersection of all classes in \mathcal{C} . We will show that if one of the properties (a)-(e) is satisfied by all classes in \mathcal{C} , then $\bigcap \mathcal{C}$ satisfies that property as well. The result requested in the problem then follows as an immediate corollary.

It is clear that if every $\mathcal{R} \in \mathcal{C}$ satisfies (a), then so does $\bigcap \mathcal{C}$. Suppose every $\mathcal{R} \in \mathcal{C}$ satisfies (b). If $A, B \in \bigcap \mathcal{C}$ then $A, B \in \mathcal{R}$ for every $\mathcal{R} \in \mathcal{C}$. Hence $A - B \in \mathcal{R}$ for all $\mathcal{R} \in \mathcal{C}$, and it follows that $A - B \in \bigcap \mathcal{C}$. The argument for (c) is similar (with $A \cup B$ in place of $A - B$), and (d) is obvious.

Finally, suppose that every $\mathcal{R} \in \mathcal{C}$ satisfies (e). If $A_1, A_2, \dots \in \bigcap \mathcal{C}$ then $A_1, A_2, \dots \in \mathcal{R}$ for every $\mathcal{R} \in \mathcal{C}$. Hence $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$ for all $\mathcal{R} \in \mathcal{C}$, and it follows that $\bigcup_{n=1}^{\infty} A_n \in \bigcap \mathcal{C}$.

1.1.5

If \mathcal{D} is any class of sets, then there exists a unique ring \mathcal{R}_0 such that (i) $\mathcal{R}_0 \supset \mathcal{D}$, and (ii) any ring \mathcal{R} containing \mathcal{D} contains also \mathcal{R}_0 . \mathcal{R}_0 is called the *ring generated by \mathcal{D}* , and is denoted by $\mathcal{R}(\mathcal{D})$.

Solution. Let \mathcal{R}_0 be the intersection of all rings containing \mathcal{D} . This is a ring by the previous exercise, and it satisfies the properties (i) and (ii). To see that it is unique, let \mathcal{R}'_0 also be a ring satisfying (i) and (ii). Then $\mathcal{R}_0 \subset \mathcal{R}'_0$ and $\mathcal{R}'_0 \subset \mathcal{R}_0$ by property (ii), so $\mathcal{R}_0 = \mathcal{R}'_0$.

1.1.6

If \mathcal{D} is any class of sets, then there exists a unique σ -ring \mathcal{S}_0 such that (i) $\mathcal{S}_0 \supset \mathcal{D}$, and (ii) any σ -ring containing \mathcal{D} contains also \mathcal{S}_0 . We call \mathcal{S}_0 the *σ -ring generated by \mathcal{D}* , and denote it by $\mathcal{S}(\mathcal{D})$. A similar result holds for σ -algebras, and we speak of the *σ -algebra generated by \mathcal{D}* .

Solution. By the same argument as in the previous exercise, \mathcal{S}_0 is the intersection of all σ -rings containing \mathcal{D} . Similarly the σ -algebra generated by \mathcal{D} is the intersection of all σ -algebras containing \mathcal{D} .

1.1.7

If \mathcal{D} is any class of sets, then every set in $\mathcal{R}(\mathcal{D})$ can be covered by (that is, is contained in) a finite union of sets of \mathcal{D} . [*Hint:* The class \mathcal{K} of sets that can be covered by finite unions of sets of \mathcal{D} forms a ring.]

Solution. Let \mathcal{K} be the class of all sets that can be covered by a finite union of sets in \mathcal{D} . Certainly $\emptyset \in \mathcal{K}$, since \emptyset is a subset of the empty union. If $A, B \in \mathcal{K}$, then

$$A \subset \bigcup_{i=1}^m E_i, \quad B \subset \bigcup_{i=1}^n F_i,$$

for some sets $E_1, \dots, E_m, F_1, \dots, F_n \in \mathcal{D}$. (Note that m or n can be zero, in which case the corresponding union is empty.) Thus

$$A - B \subset A \subset \bigcup_{i=1}^m E_i$$

and

$$A \cup B \subset \left(\bigcup_{i=1}^m E_i \right) \cup \left(\bigcup_{j=1}^n F_j \right),$$

so both $A - B$ and $A \cup B$ are elements of \mathcal{K} .

The above shows that \mathcal{K} is a ring, and certainly $\mathcal{D} \subset \mathcal{K}$. Hence $\mathcal{R}(\mathcal{D}) \subset \mathcal{K}$ by Problem 1.1.5, and it follows that every set in $\mathcal{R}(\mathcal{D})$ can be covered by a finite union of sets in \mathcal{D} .

Section 1.1 – Definition of Measure

Problems

1.2.1

If μ satisfies the properties (i)-(iii) in Definition 1.2.1, and if $\mu(E) < \infty$ for at least one set E , then (iv) is also satisfied.

Solution. We have

$$\mu(E) = \mu(E \cup \emptyset) = \mu(E) + \mu(\emptyset),$$

hence $\mu(\emptyset) = 0$.

1.2.2

Let X be an infinite space. Let \mathcal{A} be the class of all subsets of X . Define $\mu(E) = 0$ if E is finite and $\mu(E) = \infty$ if E is infinite. Then μ is finitely additive but not completely additive.

Solution. Suppose $A, B \in \mathcal{A}$. Note that $A \cup B$ is finite if both A and B are finite, but infinite otherwise. Hence

$$\mu(A \cup B) = 0 = \mu(A) + \mu(B)$$

in the former case, and

$$\mu(A \cup B) = \infty = \mu(A) + \mu(B)$$

in the latter. This proves that μ is additive; *finite* additivity follows by a simple induction argument.

Let (x_n) be a sequence of distinct points in X . Then $\bigcup_{n=1}^{\infty} \{x_n\}$ is an infinite set, so

$$\mu\left(\bigcup_{n=1}^{\infty} \{x_n\}\right) = \infty,$$

but

$$\sum_{n=1}^{\infty} \mu(\{x_n\}) = 0.$$

Thus μ is not completely additive.

1.2.3

If μ is a measure on a σ -algebra \mathcal{A} , and if E, F are sets of \mathcal{A} , then

$$\mu(E) + \mu(F) = \mu(E \cup F) + \mu(E \cap F).$$

Solution. If $\mu(F) = \infty$, then $\mu(E \cup F) = \infty$ by Theorem 1.2.1(i), and the given equality holds. If $\mu(F) < \infty$, then

$$\begin{aligned} \mu(E \cup F) &= \mu[E \cup (F - (E \cap F))] \\ &= \mu(E) + \mu[F - (E \cap F)] \\ &= \mu(E) + \mu(F) - \mu(E \cap F), \end{aligned}$$

with the last equality following from Theorem 1.2.1(ii). Note that $E \cap F \subset F$ so that $\mu(E \cap F) \leq \mu(F) < \infty$. Hence we can rearrange the above to yield

$$\mu(E) + \mu(F) = \mu(E \cup F) + \mu(E \cap F).$$

1.2.6

Give an example of a measure μ and a monotone-decreasing sequence $\{E_n\}$ of \mathcal{A} such that $\mu(E_n) = \infty$ for all n , and $\mu(\lim_n E_n) = 0$.

Solution. Let $X = \mathbb{R}$ and let $\mathcal{A} = \mathcal{P}(\mathbb{R})$ (the power set of \mathbb{R} ; this is easily seen to be a σ -algebra). Define $\mu : \mathcal{A} \rightarrow [0, \infty]$ such that $\mu(E)$ is the number of points in E (with $\mu(E) = \infty$ if E is infinite). This is easily seen to be a measure.

For each $n \in \mathbb{N}$, let $E_n = (0, 1/n)$. Then (E_n) is a monotone decreasing sequence of sets in \mathcal{A} , $\mu(E_n) = \infty$ for all n , and

$$\mu\left(\lim_n E_n\right) = \mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \mu(\emptyset) = 0.$$

Section 1.3 – Outer Measure

Problems

1.3.1

Define $\mu^*(E)$ as the number of points in E if E is finite and $\mu^*(E) = \infty$ if E is infinite. Show that μ^* is an outer measure. Determine the measurable sets.

Solution. Of the properties listed in Definition 1.3.1, only countable subadditivity is non-obvious for μ^* . But let us start with proving finite subadditivity.

Let A and B be sets. If either is infinite, then so is $A \cup B$, hence

$$\mu^*(A \cup B) = \infty = \mu^*(A) + \mu^*(B).$$

If both A and B are finite sets, then

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B - A) \leq \mu^*(A) + \mu^*(B)$$

by basic set-theoretic considerations. Thus μ^* is subadditive, and finite subadditivity follows by induction on the number of sets in the union.

Now, let (E_n) be a sequence of sets. If infinitely many of the sets E_n are nonempty, then $\sum_n \mu^*(E_n) = \infty$, and

$$\mu^*\left(\bigcup_n E_n\right) \leq \sum_n \mu^*(E_n)$$

follows. If only finitely many of the sets E_n are nonempty, let $E_{n_1}, E_{n_2}, \dots, E_{n_k}$ be those sets. Then

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) = \mu^*\left(\bigcup_{i=1}^k E_{n_i}\right) \leq \sum_{i=1}^k \mu^*(E_{n_i}) = \sum_{n=1}^{\infty} \mu^*(E_n),$$

by finite subadditivity. This proves that μ^* is countably subadditive, and hence that μ^* is an outer measure.

Note that μ^* is *additive* on disjoint sets; if $A \cap B = \emptyset$, then $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$. In particular,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A - E)$$

for all sets A, E . That is, all sets are measurable.

1.3.2

Define $\mu^*(\emptyset) = 0$, $\mu^*(E) = 1$ if $E \neq \emptyset$. Show that μ^* is an outer measure, and determine the measurable sets.

Solution. As in the previous exercise, the only slightly non-obvious property is countable subadditivity. Hence, let (E_n) be a sequence of sets. If all the sets E_n are empty, then certainly

$$\mu^*\left(\bigcup_n E_n\right) = 0 = \sum_n \mu^*(E_n).$$

If not, then there is some m such that $E_m \neq \emptyset$, and it follows that

$$\mu^*\left(\bigcup_n E_n\right) = 1 = \mu^*(E_m) \leq \sum_n \mu^*(E_n).$$

Thus μ^* is indeed countably subadditive, and therefore also an outer measure.

The empty set is measurable:

$$\mu^*(A \cap \emptyset) + \mu^*(A - \emptyset) = \mu^*(\emptyset) + \mu^*(A) = \mu^*(A)$$

for all sets A . It follows by Theorem 1.3.1 that X is measurable as well (the measurable sets make up a σ -algebra). Indeed \emptyset and X are the only measurable sets. To see this, let E be any set other than those two (this requires that X contains at least two elements). Then both E and E^c are nonempty, so

$$\mu^*(X \cap E) + \mu^*(X - E) = \mu^*(E) + \mu^*(E^c) = 2 > 1 = \mu^*(X).$$

Section 1.4 – Construction of Outer Measures

Problems

1.4.4

If \mathcal{K} is a σ -algebra and λ is a measure on \mathcal{K} , then $\mu^*(A) = \lambda(A)$ for any $A \in \mathcal{K}$. [*Hint:* $\mu^*(A) = \inf\{\lambda(E); E \in \mathcal{K}, E \supset A\}$.]

Solution. Note that the description of μ^* can be simplified when \mathcal{K} is a σ -algebra and λ is a measure. For suppose that $A \subset X$, $E_n \in \mathcal{K}$ ($n = 1, 2, \dots$), and $A \subset \bigcup_n E_n$. Then $E := \bigcup_n E_n \in \mathcal{K}$, and $\lambda(E) \leq \sum_n \lambda(E_n)$ by Theorem 1.2.2. Hence

$$\mu^*(A) = \inf\{\lambda(E); E \in \mathcal{K}, E \supset A\}.$$

Now, suppose that $A \in \mathcal{K}$. Certainly $\lambda(A)$ is an element of $\{\lambda(E); E \in \mathcal{K}, E \supset A\}$. And if $E \in \mathcal{K}$ and $E \supset A$, then $\lambda(E) \geq \lambda(A)$ by Theorem 1.2.1(i). Thus

$$\lambda(A) = \inf\{\lambda(E); E \in \mathcal{K}, E \supset A\} = \mu^*(A).$$

1.4.5

If \mathcal{K} is a σ -algebra and λ is a measure on \mathcal{K} , then every set in \mathcal{K} is μ^* -measurable.

Solution. Recall the simplified description of μ^* from the previous problem. Let $E \in \mathcal{K}$ and $A \subset X$. For every $\epsilon > 0$ there exists $F \in \mathcal{K}$ such that $F \supset A$ and

$$\mu^*(A) + \epsilon > \lambda(F);$$

else $\mu^*(A)$ would not be the greatest lower bound of $\{\lambda(E); E \in \mathcal{K}, E \supset A\}$. Moreover,

$$\lambda(F) = \lambda(F \cap E) + \lambda(F - E)$$

since λ is a measure on \mathcal{K} ,

$$\lambda(F \cap E) + \lambda(F - E) = \mu^*(F \cap E) + \mu^*(F - E)$$

by what we found in the previous exercise, and finally

$$\mu^*(F \cap E) + \mu^*(F - E) \geq \mu^*(A \cap E) + \mu^*(A - E)$$

by monotonicity of the outer measure μ^* . Putting all of this together, we have

$$\mu^*(A) + \epsilon > \mu^*(A \cap E) + \mu^*(A - E)$$

for all $\epsilon > 0$, and thus

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A - E).$$

It follows that every set $E \in \mathcal{K}$ is μ^* -measurable.

Section 1.6 – The Lebesgue and the Lebesgue-Stieltjes Measures

Problems

1.6.3

The outer Lebesgue measure of a closed bounded interval $[a, b]$ on the real line is equal to $b - a$. [*Hint:* Use the Heine-Borel theorem to replace a countable covering by a finite covering.]

Solution. Suppose (E_n) is a sequence of elements of \mathcal{K} (i.e. a sequence of open intervals) such that $[a, b] \subset \bigcup_{n=1}^{\infty} E_n$. The collection $\{E_n\}$ constitutes an *open cover* of $[a, b]$. By the Heine-Borel theorem $[a, b]$ is compact, hence there exists a *finite subcover* $\{E_{n_1}, \dots, E_{n_k}\}$, such that $[a, b] \subset \bigcup_{i=1}^k E_{n_i}$.

Assume without loss of generality that $E_{n_i} \cap [a, b] \neq \emptyset$ for all i ; otherwise we can simply remove those E_{n_i} that are disjoint with $[a, b]$ and still have a finite subcover. Write $E_{n_i} = (a_i, b_i)$ for each i , and define

$$\alpha = \min\{a_1, \dots, a_k\}, \quad \beta = \max\{b_1, \dots, b_k\}.$$

It is clear that α and β are the infimum and supremum, respectively, of $\bigcup_{i=1}^k E_{n_i}$. Note that $\alpha = a_j$ for some j , and $a_j < a < b_j$ since E_{n_j} and $[a, b]$ have nonempty intersection. Thus $(\alpha, a] \subset \bigcup_{i=1}^k E_{n_i}$, and similarly $[b, \beta) \subset \bigcup_{i=1}^k E_{n_i}$. It follows that

$$\bigcup_{i=1}^k E_{n_i} = (\alpha, \beta) \in \mathcal{K}.$$

Finally note that λ is finitely subadditive. (This is easily proven with induction.) (TODO: This is not convincing; use better proof from Rosenthal notes.) Thus,

$$\sum_{n=1}^{\infty} \lambda(E_n) \geq \sum_{i=1}^k \lambda(E_{n_i}) \geq \lambda[(\alpha, \beta)] = \beta - \alpha > b - a.$$

It follows that $b - a$ is a lower bound of the set

$$\Lambda([a, b]) := \left\{ \sum_{n=1}^{\infty} \lambda(E_n); E_n \in \mathcal{K}, \bigcup_{n=1}^{\infty} E_n \supset [a, b] \right\}.$$

Moreover, for every $\epsilon > 0$ we have

$$[a, b] \subset \left(a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2}\right) \in \mathcal{K}$$

and

$$\lambda \left[\left(a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2}\right) \right] = b - a + \epsilon.$$

Hence $b - a$ is the *greatest* lower bound of $\Lambda([a, b])$, and $\mu^*([a, b]) = b - a$.

1.6.4

The outer Lebesgue measure of each of the intervals (a, b) , $[a, b)$, $(a, b]$ is equal to $b - a$.

Solution. Recall that μ^* is monotone, on account of being an outer measure. Hence $\mu^*[(a, b)] \leq \mu^*([a, b]) = b - a$, the latter equality being the result of the previous problem. Moreover, for all $\epsilon \in (0, b - a)$ we have

$$\left(a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2}\right) \subset (a, b),$$

so that

$$\mu^*[(a, b)] \geq \mu^* \left[\left(a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2}\right) \right] = b - a + \epsilon.$$

Thus $\mu^*[(a, b)] \geq b - a$, and it follows that $\mu^*[(a, b)] = b - a$.

The outer measures of $[a, b)$ and $(a, b]$ follow immediately by monotonicity:

$$\mu^*[(a, b)] \leq \mu^*([a, b)) \leq \mu^*([a, b]),$$

so that $\mu^*([a, b)) = b - a$. Similarly for $(a, b]$.

1.6.5

Consider the transformation $Tx = \alpha x + \beta$ from the real line onto itself, where α, β are real numbers and $\alpha \neq 0$. It maps sets E onto sets $T(E)$. Denote by μ (μ^*) the Lebesgue measure (outer measure) on the real line. Prove

- (a) For any set E , $\mu^*(T(E)) = |\alpha|\mu^*(E)$.
- (b) E is Lebesgue-measurable if and only if $T(E)$ is Lebesgue-measurable.
- (c) If E is Lebesgue-measurable, then $\mu(T(E)) = |\alpha|\mu(E)$.

Solution. Let us start with a couple of simple observations:

- T is bijective, with inverse given by

$$T^{-1}(x) = \frac{x - \beta}{\alpha}.$$

- Suppose $I = (a, b)$. Then

$$T(I) = (\alpha a + \beta, \alpha b + \beta)$$

if $\alpha > 0$, and

$$T(I) = (\beta b + \beta, \alpha a + \beta)$$

if $\alpha < 0$. Either way,

$$\mu^*[T(I)] = |\alpha|(b - a) = |\alpha|\mu^*(I),$$

where we have used one of the results of the previous exercise. Similarly, $T^{-1}(I)$ is an open interval and

$$\mu^*[T^{-1}(I)] = |\alpha|^{-1}\mu^*(I).$$

Of course, the latter two identities still hold if $I = \emptyset$. Hence they hold for all $I \in \mathcal{K}$.

Also, let us use the notation

$$\Lambda(E) = \left\{ \sum_{n=1}^{\infty} \lambda(I_n); I_n \in \mathcal{K}, \bigcup_{n=1}^{\infty} I_n \supset E \right\}$$

for all $E \subset \mathbb{R}$.

- (a) Suppose (I_n) is a sequence in \mathcal{K} (i.e. a sequence of open intervals) and $E \subset \bigcup_n I_n$. Then $T(I_n) \in \mathcal{K}$ for every n ,

$$T(E) \subset T\left(\bigcup_n I_n\right) = \bigcup_n T(I_n),$$

and

$$\sum_n \lambda[T(I_n)] = |\alpha| \sum_n \lambda(I_n).$$

Thus, if $s \in \Lambda(E)$, then $|\alpha|s \in \Lambda[T(E)]$. It follows that

$$\mu^*[T(E)] = \inf \Lambda[T(E)] \leq |\alpha| \inf \Lambda(E) = |\alpha| \mu^*(E).$$

Conversely, suppose (J_n) is a sequence in \mathcal{K} and $T(E) \subset \bigcup_n J_n$. Then $T^{-1}(J_n) \in \mathcal{K}$ for all n ,

$$E = T^{-1}[T(E)] \subset T^{-1}\left(\bigcup_n J_n\right) = \bigcup_n T^{-1}(J_n),$$

and

$$\sum_n \lambda[T^{-1}(J_n)] = |\alpha|^{-1} \sum_n \lambda(J_n).$$

Hence, by the same logic as above, we find that $\mu^*(E) \leq |\alpha|^{-1} \mu^*[T(E)]$, and it follows that

$$\mu^*[T(E)] = |\alpha| \mu^*(E).$$

- (b) Note that if $f : X \rightarrow Y$ is a bijective function (between arbitrary sets X, Y), then

$$\begin{aligned} f^{-1}[f(A)] &= A, \\ f(A \cup B) &= f(A) \cup f(B), \\ f(A - B) &= f(A) - f(B), \\ f[f^{-1}(C)] &= C, \end{aligned}$$

for all $A, B \subset X$ and $C \subset Y$.

Suppose that E is measurable:

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A - E)$$

for all $A \subset \mathbb{R}$. Then, for all $B \subset \mathbb{R}$, we have

$$\begin{aligned} \mu^*[B \cap T(E)] + \mu^*[B - T(E)] &= \mu^*[T(T^{-1}(B) \cap E)] + \mu^*[T(T^{-1}(B) - E)] \\ &= |\alpha| \mu^*[T^{-1}(B) \cap E] + |\alpha| \mu^*[T^{-1}(B) - E] \\ &= |\alpha| \mu^*[T^{-1}(B)] \\ &= \mu^*(B), \end{aligned}$$

so that $T(E)$ is measurable.

Conversely, suppose that $T(E)$ is measurable. Then, for all $A \subset \mathbb{R}$,

$$\begin{aligned} \mu^*(A \cap E) + \mu^*(A - E) &= \mu^*[T^{-1}(T(A) \cap T(E))] + \mu^*[T^{-1}(T(A) - T(E))] \\ &= |\alpha|^{-1} \mu^*[T(A) \cap T(E)] + |\alpha|^{-1} \mu^*[T(A) - T(E)] \\ &= |\alpha|^{-1} \mu^*[T(A)] \\ &= \mu^*(A), \end{aligned}$$

so that E is measurable.

- (c) This is immediate given (a), (b), and the definition of the Lebesgue-measure. First, $T(E)$ is Lebesgue-measurable by (b). Next, $\mu(E) = \mu^*(E)$ and $\mu[T(E)] = \mu^*[T(E)]$ since μ is simply the restriction of μ^* to the measurable sets. Finally, $\mu^*[T(E)] = |\alpha|\mu^*(E)$ by (a).

Chapter 2 – Integration

Section 2.1 – Definition of Measurable Functions

Problems

2.1.6

The *characteristic function* of a set E is the function χ_E defined by

$$\chi_E(x) = \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$$

Prove that the set E is measurable if and only if the function χ_E is measurable.

Solution. Suppose $E \in \mathcal{A}$. For all $c \in \mathbb{R}$,

$$\chi_E^{-1}\{(-\infty, c)\} = \{x \in X; \chi_E(x) < c\} = \begin{cases} \emptyset & (c \leq 0), \\ E^c & (0 < c \leq 1), \\ X & (c > 1), \end{cases}$$

so that $\chi_E^{-1}\{(-\infty, c)\} \in \mathcal{A}$. By Theorem 2.1.1, χ_E is measurable.

Conversely, suppose χ_E is measurable. Then E is measurable, since

$$E = X - E^c = \chi^{-1}\{(-\infty, 2)\} - \chi^{-1}\{(-\infty, 1)\}.$$

2.1.9

If f is measurable, then $|f|$ and $|f|^2$ are measurable.

Solution. If $c \leq 0$, then

$$(|f|)^{-1}\{(-\infty, c)\} = (|f|^2)^{-1}\{(-\infty, c)\} = \emptyset \in \mathcal{A},$$

since $|f|$ and $|f|^2$ are nonnegative functions.

Let $c > 0$. Then

$$(|f|)^{-1}\{(-\infty, c)\} = \{x \in X; -c < f(x) < c\} = f^{-1}\{(-c, c)\}.$$

The set $(-c, c)$ is open, hence $f^{-1}\{(-c, c)\} \in \mathcal{A}$ by the measurability of f . Similarly,

$$(|f|^2)^{-1}\{(-\infty, c)\} = f^{-1}\{(-\sqrt{c}, \sqrt{c})\} \in \mathcal{A}.$$

Finally,

$$(|f|)^{-1}\{+\infty\} = (|f|^2)^{-1}\{+\infty\} = f^{-1}\{+\infty\} \cup f^{-1}\{-\infty\} \in \mathcal{A}$$

by the measurability of f , and

$$(|f|)^{-1}\{-\infty\} = (|f|^2)^{-1}\{-\infty\} = \emptyset \in \mathcal{A}$$

since $|f|$ and $|f|^2$ are nonnegative. Thus, both $|f|$ and $|f|^2$ are measurable by Theorem 2.1.1.

2.1.10

A monotone function defined on the real line is Lebesgue-measurable.

Solution. Let f be a monotone increasing extended real-valued function on \mathbb{R} ;

$$(\forall x, y \in \mathbb{R}) : \quad x < y \implies f(x) \leq f(y).$$

Given any $c \in \mathbb{R}$, let

$$\xi_c = \inf\{x \in X; f(x) \geq c\}.$$

We need to consider two cases: $f(\xi_c) < c$ and $f(\xi_c) \geq c$. In the former case, $f(x) < c$ for all $x \leq \xi_c$ and $f(x) \geq c$ for all $x > \xi_c$ (by monotonicity). Hence

$$f^{-1}\{(-\infty, c)\} = (-\infty, \xi_c].$$

This is a Borel set, hence also a Lebesgue set (see Problem 1.9.3). In the latter case, $f(x) < c$ for all $x < \xi_c$ and $f(x) \geq c$ for all $x \geq \xi_c$, so that

$$f^{-1}\{(-\infty, c)\} = (-\infty, \xi_c),$$

which is Lebesgue-measurable. Since c was arbitrary, we conclude that f is measurable, by Theorem 2.1.1.

The proof for f monotone decreasing is similar.

Section 2.2 – Operations on Measurable Functions

Problems

2.2.2

Let $g(u_1, \dots, u_k)$ be a continuous function in \mathbb{R}^k , and let $\varphi_1, \dots, \varphi_k$ be measurable functions. Prove that the composite function $h(x) = g[\varphi_1(x), \dots, \varphi_k(x)]$ is a measurable function. Note that as a special case we may conclude that

$$\max(\varphi, \dots, \varphi_n) \quad \text{and} \quad \min(\varphi, \dots, \varphi_n)$$

are measurable functions.

Solution. We will use the following fact, which may be proven in a course in topology:

\mathbb{R}^k has a countable basis of product open subsets. Hence, if U is an open subset of \mathbb{R}^k , then there are open subsets $U_{ni} \subset \mathbb{R}$ for $n = 1, 2, \dots$ and $i = 1, \dots, k$ such that

$$U = \bigcup_{n=1}^{\infty} (U_{n1} \times \cdots \times U_{nk}).$$

We are assuming that g is real-valued, likewise for the functions φ_i . Let $c \in \mathbb{R}$. Note that $g^{-1}\{(-\infty, c)\}$ is open by continuity of g . Thus

$$g^{-1}\{(-\infty, c)\} = \bigcup_{n=1}^{\infty} (U_{n1} \times \cdots \times U_{nk})$$

for some open subsets $U_{ni} \subset \mathbb{R}$. Hence

$$\begin{aligned} h^{-1}\{(-\infty, c)\} &= \{x \in X; g(\varphi_1(x), \dots, \varphi_k(x)) \leq c\} \\ &= \{x \in X; (\varphi_1(x), \dots, \varphi_k(x)) \in g^{-1}\{(-\infty, c)\}\} \\ &= \bigcup_{n=1}^{\infty} \{x \in X; (\varphi_1(x), \dots, \varphi_k(x)) \in U_{n1} \times \cdots \times U_{nk}\} \\ &= \bigcup_{n=1}^{\infty} \bigcap_{i=1}^k \{x \in X; \varphi_i(x) \in U_{ni}\} \\ &= \bigcup_{n=1}^{\infty} \bigcap_{i=1}^k \varphi_i^{-1}(U_{ni}). \end{aligned}$$

The sets $\varphi_i(U_{ni})$ are measurable since the functions φ_i are measurable. It follows that $h^{-1}\{(-\infty, c)\}$ is measurable, and thus that h is measurable, by Theorem 2.1.1.

To apply the above to the max and min functions $\mathbb{R}^k \rightarrow \mathbb{R}$ we must show that they are continuous. Let $a < b$ and note that

$$\begin{aligned} \max^{-1}\{(a, b)\} &= \{(x_1, \dots, x_k) \in \mathbb{R}^k; x_i > a \text{ for some } i\} \\ &\cap \{(x_1, \dots, x_k) \in \mathbb{R}^k; x_i < b \text{ for all } i\}. \end{aligned}$$

Both sets in the above binary intersection are easily seen to be open by considering ϵ -neighborhoods about their points. It follows that $\max^{-1}(U)$ is open for all open subsets $U \in \mathbb{R}^k$, since every such U can be written as a countable union of open intervals. Thus max is continuous, and one similarly shows that min is continuous.

2.2.3

Let $f(x)$ be a measurable function and define

$$g(x) = \begin{cases} \frac{1}{f(x)}, & \text{if } f(x) \neq 0, \\ 0, & \text{if } f(x) = 0. \end{cases}$$

Prove that g is measurable.

Solution. For $c < 0$,

$$g^{-1}\{(-\infty, c)\} = \{x; 1/f(x) < c\} = \{x; 1/c < f < 0\} = f^{-1}\{(1/c, 0)\},$$

which is measurable by the measurability of f . Next,

$$g^{-1}\{(-\infty, 0)\} = \{x; 1/f(x) < 0\} = \{x; f(x) < 0\} = f^{-1}\{(-\infty, 0)\},$$

also measurable. Note that if we take the natural convention (unfortunately not addressed in the text) that $x/(\pm\infty) = 0$ for all $x \in \mathbb{R}$, then

$$g^{-1}\{0\} = \{x; f(x) = 0\} \cup \{x; f(x) = \pm\infty\} = f^{-1}\{0\} \cup f^{-1}\{\pm\infty\}.$$

Hence, for $c > 0$,

$$\begin{aligned} g^{-1}\{(0, c)\} &= g^{-1}\{(-\infty, 0)\} \cup g^{-1}\{0\} \cup g^{-1}\{(0, \infty)\} \\ &= f^{-1}\{(-\infty, 0)\} \cup f^{-1}\{0\} \cup f^{-1}\{\pm\infty\} \cup f^{-1}\{(1/c, \infty)\} \\ &= f^{-1}\{(-\infty, 0]\} \cup f^{-1}\{\pm\infty\} \cup f^{-1}\{(1/c, \infty)\}, \end{aligned}$$

which is measurable by the measurability of f (see Problem 2.1.4). Finally, $g^{-1}\{\pm\infty\} = \emptyset$, and it follows by Theorem 2.1.1 that g is measurable.

Section 2.3 – Egoroff's Theorem

Problems

2.3.2

Let $\{f_n\}$ be a sequence of measurable functions in a finite measure space X . Suppose that for almost every x , $\{f_n(x)\}$ is a bounded set. Then for any $\epsilon > 0$ there exist a positive number c and a measurable set E with $\mu(X - E) < \epsilon$, such that $|f_n(x)| \leq c$ for all $x \in E$, $n = 1, 2, \dots$.

Solution. The definition we have for 'bounded set' applies to metric spaces, and it does not make much sense here since the functions f_n may be extended real-valued. Hence we will assume that ' $\{f_n(x)\}$ is a bounded set' means that $\sup_n |f_n(x)| < \infty$.

Let $g = \sup_n |f_n|$, and note that g is measurable by Problem 2.1.9 and Theorem 2.2.3. Let $F = \{x; g(x) < \infty\}$. Notice that $g(x) < \infty$ if and only if $\{f_n(x)\}$ is bounded. Hence $\mu(X - F) = 0$.

For $k = 1, 2, \dots$, define $F_k = \{x; g(x) \leq k\}$. Then $F_1 \subset F_2 \subset \dots$ and $\lim_k F_k = \bigcup_{k=1}^{\infty} F_k = F$. By Theorem 1.2.1(iv),

$$\lim_k \mu(X - F_k) = \mu(X - F) = 0.$$

Given any $\epsilon > 0$, there exists a positive integer K such that $\mu(X - F_k) < \epsilon$ for all $k \geq K$. In particular $\mu(X - F_K) < \epsilon$, and $g(x) \leq K$ for all $x \in F_K$, which means that $|f_n(x)| \leq K$ for all $x \in F_K$.

Section 2.4 – Convergence in Measure

Problems

2.4.3

Prove the following result (which immediately yields another proof of Corollary 2.4.2): Let f_n ($n = 1, 2, \dots$) and f be a.e. real-valued measurable functions in a finite measure space. For any $\epsilon > 0$, $n \geq 1$, let

$$E_n(\epsilon) = \{x; |f_n(x) - f(x)| \geq \epsilon\}.$$

Then $\{f_n\}$ converges a.e. to f if and only if

$$\lim_{n \rightarrow \infty} \mu \left[\bigcup_{m=n}^{\infty} E_m(\epsilon) \right] = 0 \quad \text{for any } \epsilon > 0. \quad (2.4.2)$$

[Hint: Let $F = \{x; \{f_n(x)\} \text{ is not convergent to } f(x)\}$. Then $F = \bigcup_{k=1}^{\infty} \overline{\lim_n E_n(1/k)}$. Show that $\mu(F) = 0$ if and only if (2.4.2) holds.]

Solution. Define

$$F = \bigcup_{k=1}^{\infty} \overline{\lim_n E_n \left(\frac{1}{k} \right)} = \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m \left(\frac{1}{k} \right).$$

Note that

$$\begin{aligned} x \in F &\iff \exists k, \forall n, \exists m \geq n, |f_m(x) - f(x)| \geq \frac{1}{k} \\ &\iff \neg \left(\forall k, \exists n, \forall m \geq n, |f_m(x) - f(x)| < \frac{1}{k} \right) \\ &\iff f_n(x) \not\rightarrow f(x), \end{aligned}$$

so that

$$F = \{x; f_n(x) \not\rightarrow f(x)\}.$$

Suppose (2.4.2) holds. Fix $\delta > 0$. For every positive integer k , there exists a positive integer n_k such that $n \geq n_k$ implies

$$\mu \left[\bigcup_{m=n}^{\infty} E_m \left(\frac{1}{k} \right) \right] < \frac{\delta}{2^k}.$$

By subadditivity and monotonicity,

$$\begin{aligned}\mu(F) &= \mu \left[\bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m \left(\frac{1}{k} \right) \right] \leq \sum_{k=1}^{\infty} \mu \left[\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m \left(\frac{1}{k} \right) \right] \\ &\leq \sum_{k=1}^{\infty} \mu \left[\bigcup_{m=n_k}^{\infty} E_m \left(\frac{1}{k} \right) \right] < \sum_{k=1}^{\infty} \frac{\delta}{2^k} = \delta.\end{aligned}$$

Since δ was arbitrary, $\mu(F) = 0$, and it follows that $f_n \rightarrow f$ a.e.

Conversely, suppose $f_n \rightarrow f$ a.e., so that $\mu(F) = 0$. By monotonicity and Theorem 1.2.2,

$$0 = \mu(F) = \mu \left[\bigcup_{k=1}^{\infty} \overline{\lim}_n E_n \left(\frac{1}{k} \right) \right] \geq \mu \left[\overline{\lim}_n E_n \left(\frac{1}{l} \right) \right] \geq \overline{\lim}_n \mu \left[E_n \left(\frac{1}{l} \right) \right]$$

for all positive integers l . But of course $\overline{\lim}_n \mu [E_n (1/l)] \geq \underline{\lim}_n \mu [E_n (1/l)] \geq 0$ since μ is nonnegative, so $\lim_n \mu [E_n (1/l)]$ exists and is equal to zero. Note that the sets $\bigcup_{m=n}^{\infty} E_m (1/l)$ are decreasing, so their limit as $n \rightarrow \infty$ exists. Hence we can apply Corollary 1.2.3 and monotonicity to find that

$$\begin{aligned}\lim_n \mu \left[\bigcup_{m=n}^{\infty} E_m \left(\frac{1}{l} \right) \right] &= \mu \left[\lim_n \bigcup_{m=n}^{\infty} E_m \left(\frac{1}{l} \right) \right] = \mu \left[\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m \left(\frac{1}{l} \right) \right] \\ &\leq \mu \left[\bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m \left(\frac{1}{k} \right) \right] = \mu(F) = 0.\end{aligned}$$

Finally, given $\epsilon > 0$, note that

$$E_n(\epsilon) \subset E_n \left(\frac{1}{\lceil 1/\epsilon \rceil} \right).$$

Hence

$$\lim_n \mu \left[\bigcup_{m=n}^{\infty} E_m(\epsilon) \right] \leq \lim_n \mu \left[\bigcup_{m=n}^{\infty} E_m \left(\frac{1}{\lceil 1/\epsilon \rceil} \right) \right] \leq 0$$

by monotonicity, and (2.4.2) follows.

2.4.4

Let X be the set of all positive integers, \mathcal{A} the class of all subsets of X , and $\mu(E)$ (for any $E \in \mathcal{A}$) the number of points in E . Prove that in this measure space, convergence in measure is equivalent to uniform convergence.

Solution. Uniform convergence always implies convergence in measure. Conversely, suppose (f_n) converges in measure to f . Given any $\epsilon > 0$, there exists a positive integer N such that $n \geq N$ implies

$$\mu [\{x; |f_n(x) - f(x)| \geq \epsilon\}] < 1.$$

That is, for $n \geq N$ the set $\{x; |f_n(x) - f(x)| \geq \epsilon\}$ is empty, which in particular means that $\sup_x |f_n(x) - f(x)| \leq \epsilon$. It follows that $f_n \rightarrow f$ uniformly.

Section 2.5 – Integrals of Simple Functions

Problems

2.5.2

An integrable simple function f is equal a.e. to zero if and only if $\int_E f d\mu = 0$ for any measurable set E .

Solution. Let f be an integrable simple function. Then f can be written in the form

$$f = \sum_{i=1}^n \alpha_i \chi_{E_i},$$

for mutually disjoint sets E_1, \dots, E_n , with all $\alpha_i \neq 0$, and all $\mu(E_i) < \infty$.

Suppose $f = 0$ a.e., and let E be any measurable set. By Theorem 2.5.1(b) and (g),

$$0 \leq \int_E f d\mu \leq \int f d\mu = \sum_{i=1}^n \alpha_i \mu(E_i).$$

But $\mu(E_i) = 0$ since $f = 0$ a.e., so $\int_E f d\mu = 0$.

Conversely, suppose $\int_E f d\mu = 0$ for all measurable sets E . Then

$$\alpha_i \mu(E_i) = \int_{E_i} f d\mu = 0,$$

so that $\mu(E_i) = 0$, for all $i \in \{1, \dots, n\}$. It follows that $f = 0$ a.e.

Section 2.6 – Definition of the Integral

Problems

2.6.3

Let f be a measurable function. Prove that f is integrable if and only if f^+ and f^- are integrable, or if and only if $|f|$ is integrable.

Solution. Let f be measurable. We must prove the equivalence of the following statements:

- (i) f is integrable.
- (ii) f^+ and f^- are integrable.
- (iii) $|f|$ is integrable.

We will first show that (iii) \implies (ii), then that (ii) \implies (i), and finally that (i) \implies (iii).

Suppose that $|f|$ is integrable. Let $E = \{x; f(x) \geq 0\} = f^{-1}[0, \infty)$, and note that E is measurable since f is. There exists a Cauchy in the mean sequence (g_n) of integrable simple functions converging to $|f|$ a.e., and the sequence $(\chi_E g_n)$

is easily seen to satisfy the corresponding properties with respect to f^+ . Since f^+ is measurable by Problem 2.1.8, this implies that it is integrable. The proof that f^- is integrable is similar.

Next, suppose that f^+ and f^- are integrable. Then there exist Cauchy in the mean sequences (g_n) and (h_n) of integrable simple functions converging a.e. to f^+ and f^- , respectively. Define a new sequence (f_n) of integrable simple functions by $f_n = g_n - h_n$. Then (f_n) is Cauchy in the mean, since

$$|f_n - f_m| = |g_n - h_n - g_m + h_m| \leq |g_n - g_m| + |h_n - h_m|.$$

It also converges to f a.e. since

$$|f_n - f| = |g_n - h_n - f^+ + f^-| \leq |g_n - f^+| + |h_n - f^-|.$$

It follows that f is integrable.

Finally, assume that f is integrable. There is a Cauchy in the mean sequence (f_n) of integrable simple functions converging to f a.e. The sequence $(|f_n|)$ consists of integrable simple functions. It is Cauchy in the mean since

$$||f_n| - |f_m|| \leq |f_n - f_m|,$$

and it converges to $|f|$ a.e. since

$$||f_n| - |f|| \leq |f_n - f|.$$

Since $|f|$ is measurable by Problem 2.1.9, it follows that $|f|$ is integrable.

2.6.4

Let X be the measure space described in Problem 2.4.4. Then f is integrable if and only if the series $\sum_{n=1}^{\infty} |f(n)|$ is convergent. If f is integrable, then

$$\int f d\mu = \sum_{n=1}^{\infty} f(n).$$

Solution. Suppose f is integrable. Then there is a Cauchy in the mean sequence (f_n) of integrable simple functions converging to f a.e. We saw in the previous problem that this implies that $|f|$ is integrable, and that $(|f_n|)$ is a Cauchy in the mean sequence of integrable simple functions converging to $|f|$ a.e. Note that in this particular space convergence a.e. is the same as convergence everywhere (since the only subset with measure zero is \emptyset).

By Theorem 2.5.1(h),

$$\int |f_n| d\mu = \sum_{i=1}^{\infty} \int_{\{i\}} |f_n| d\mu = \sum_{i=1}^{\infty} |f_n(i)|.$$

Hence, in particular,

$$\int |f| d\mu = \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} |f_n(i)|.$$

Given any positive integer m , there exists n' such that

$$|f(i) - f_{n'}(i)| < 1/m \quad (i = 1, 2, \dots, m)$$

(since $f_n \rightarrow f$) and

$$\left| \sum_{i=1}^{\infty} |f_{n'}(i)| - \int |f| d\mu \right| < 1$$

(since $\sum_i |f_n(i)| \rightarrow \int |f| d\mu$). Thus

$$\sum_{i=1}^m |f(i)| \leq \sum_{i=1}^m |f(i) - f_{n'}(i)| + \sum_{i=1}^m |f_{n'}(i)| < 1 + \sum_{i=1}^{\infty} |f_{n'}(i)| < 2 + \int |f| d\mu,$$

and it follows that the series $\sum_{i=1}^{\infty} |f(i)|$ converges (to a finite number).

Conversely, assume that the series $\sum_{i=1}^{\infty} |f(i)|$ converges. Define a sequence of integrable simple functions (g_n) by

$$g_n = \sum_{i=1}^n f(i) \chi_{\{i\}}.$$

It is clear that $g_n \rightarrow f$ everywhere. Moreover, if $m > n$, then

$$\int |g_m - g_n| d\mu = \int \left| \sum_{i=n+1}^m f(i) \chi_{\{i\}} \right| d\mu = \sum_{i=n+1}^m |f(i)| \leq \sum_{i=n+1}^{\infty} |f(i)|.$$

The right-hand side goes to zero as $n \rightarrow \infty$ since $\sum_{i=1}^{\infty} |f(i)|$ is convergent, which means that $\int |g_m - g_n| d\mu \rightarrow 0$ as $n, m \rightarrow \infty$; i.e., (g_n) is Cauchy in the mean. It follows that f is integrable, with

$$\int f d\mu = \lim_{n \rightarrow \infty} \int g_n d\mu = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(i) = \sum_{i=1}^{\infty} f(i).$$

Chapter 3 – Metric Spaces

Section 3.1 – Topological and Metric Spaces

Problems

3.1.1

Prove that if (X, ρ) is a metric space, and if

$$\hat{\rho}(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)},$$

then also $(X, \hat{\rho})$ is a metric space. [*Hint:* Cf. the proof of (3.1.3).]

Solution. The only nonobvious property is the triangle inequality. Let x, y, z be arbitrary points of X . Since $t \mapsto t/(1+t)$ is monotone increasing on $[0, \infty)$, and since $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$, we have

$$\frac{\rho(x, z)}{1 + \rho(x, z)} \leq \frac{\rho(x, y) + \rho(y, z)}{1 + \rho(x, y) + \rho(y, z)}.$$

Moreover, by equation (3.1.3),

$$\frac{\rho(x, y) + \rho(y, z)}{1 + \rho(x, y) + \rho(y, z)} \leq \frac{\rho(x, y)}{1 + \rho(x, y)} + \frac{\rho(y, z)}{1 + \rho(y, z)}.$$

It follows that $\hat{\rho}(x, z) \leq \hat{\rho}(x, y) + \rho(y, z)$.

3.1.2

Let $X, \rho, \hat{\rho}$ be as in Problem 3.1.1. Prove that $\rho(x_n, x) \rightarrow 0$ if and only if $\hat{\rho}(x_n, x) \rightarrow 0$. Give an example showing that ρ and $\hat{\rho}$ are not equivalent in general.

Solution. It is clear that $\rho(x_n, x) \rightarrow 0$ implies $\hat{\rho}(x_n, x) \rightarrow 0$, since $\hat{\rho} \leq \rho$.

Conversely, suppose $\hat{\rho}(x_n, x) \rightarrow 0$. Given any $\epsilon > 0$, there is a positive integer N such that

$$\hat{\rho}(x_n, x) < \frac{\epsilon}{1 + \epsilon} \quad (n \geq N).$$

Substituting the definition of $\hat{\rho}$ and rearranging yields $\rho(x_n, x) < \epsilon$.

If ρ and $\hat{\rho}$ are equivalent then, in particular, there exists a positive constant β such that

$$\frac{\rho(x, y)}{\hat{\rho}(x, y)} \leq \beta$$

whenever $x \neq y$. But

$$\frac{\rho(x, y)}{\hat{\rho}(x, y)} = 1 + \rho(x, y),$$

so this is impossible if X is unbounded (w.r.t. ρ), say if $X = \mathbb{R}^n$ and ρ is the Euclidean metric.

3.1.6

Prove that the spaces l^1, s, c, c_0 are separable metric spaces.

Solution. For each of the spaces X we will take an arbitrary element $x \in X$ and demonstrate that every ϵ -ball around x contains a point y of a certain countable subset $Y \subset X$. It will then follow that Y is dense in X , and hence that X is separable.

Let $x = (x_i) \in l^1$. Fix $\epsilon > 0$. Since $\sum_i |x_i| < \infty$, there exists n such that

$$\sum_{i=n+1}^{\infty} |x_i| < \frac{\epsilon}{2}.$$

For $i = 1, \dots, n$, choose $y_i \in \mathbb{Q}$ (or y_i with rational real and imaginary parts in the complex case) such that $|x_i - y_i| < \epsilon/2n$, and let $y = (y_1, \dots, y_n, 0, 0, \dots)$. Then

$$\rho(x, y) = \sum_{i=1}^n |x_i - y_i| + \sum_{i=n+1}^{\infty} |x_i| < \epsilon.$$

Moreover, y is an element of the subset of l^1 consisting of sequences with rational components, with only finitely many being nonzero. This subset is easily seen to be countable, and it follows that l^1 is separable.

Next, let $x = (x_i) \in s$, and fix $\epsilon > 0$. Choose n such that

$$\sum_{i=n+1}^{\infty} \frac{1}{2^i} \frac{|x_i|}{1 + |x_i|} < \frac{\epsilon}{2}.$$

For $i = 1, \dots, n$, choose $y_i \in \mathbb{Q}$ such that $|x_i - y_i| < \epsilon/2$, and let $y = (y_1, \dots, y_n, 0, 0, \dots)$. Then

$$\rho(x, y) = \sum_{i=1}^n \frac{1}{2^i} \frac{|x_i - y_i|}{1 + |x_i - y_i|} + \sum_{i=n+1}^{\infty} \frac{1}{2^i} \frac{|x_i|}{1 + |x_i|} < \epsilon.$$

Similarly to the above, it follows that s is dense.

Finally, let $x = (x_i) \in c$. (This argument will also cover c_0 .) Let $\xi = \lim_i x_i$ and fix $\epsilon > 0$. Choose n such that $|x_i - \xi| < \epsilon/2$ for all $i \geq n$. For $i = 1, \dots, n-1$,

choose $y_i \in \mathbb{Q}$ such that $|x_i - y_i| < \epsilon$. Also choose $\eta \in \mathbb{Q}$ such that $|\xi - \eta| < \epsilon/2$ (take $\eta = \xi = 0$ in the c_0 case), so that

$$|x_i - \eta| \leq |x_i - \xi| + |\xi - \eta| < \epsilon$$

for all $i \geq n$. Let $y = (y_1, \dots, y_{n-1}, \eta, \eta, \dots)$. Then

$$\rho(x, y) = \sup_i |x_i - y_i| \leq \epsilon.$$

It follows that c is separable (and c_0 as well).

3.1.10

If $\rho(x_n, x) \rightarrow 0$, $\rho(y_n, y) \rightarrow 0$, then $\rho(x_n, y_n) \rightarrow \rho(x, y)$.

Solution. The triangle inequality yields

$$\rho(x_n, y_n) \leq \rho(x_n, x) + \rho(x, y) + \rho(y, y_n)$$

and

$$\rho(x, y) \leq \rho(x, x_n) + \rho(x_n, y_n) + \rho(y_n, y).$$

Hence

$$|\rho(x_n, y_n) - \rho(x, y)| \leq \rho(x_n, x) + \rho(y_n, y),$$

and the result follows.

Section 3.2 – L^p Spaces

Problems

3.2.4

Prove that l^p is separable if $1 \leq p < \infty$.

Solution. The argument is of the same sort as in Problem 3.1.6. Given $x = (x_i) \in l^p$ ($1 \leq p < \infty$) and $\epsilon > 0$, choose n such that

$$\sum_{i=n+1}^{\infty} |x_i|^p < \frac{\epsilon^p}{2}.$$

(This is possible since $\sum_i |x_i|^p < \infty$ for $x \in l^p$.) For $i = 1, \dots, n$, choose $y_i \in \mathbb{Q}$ such that $|x_i - y_i|^p < \epsilon^p/2n$ and let $y = (y_1, \dots, y_n, 0, 0, \dots)$. Then $\|x - y\|_p < \epsilon$, and the conclusion follows as in Problem 3.1.6.

3.2.5

Prove that l^p is not a metric space if $0 < p < 1$.

Solution. Let $x = (1, 0, 0, \dots)$, $y = (0, 0, \dots)$, $z = (0, 1, 0, 0, \dots)$. Then

$$\|x - z\|_p = 2^{1/p}$$

while

$$\|x - y\|_p + \|y - z\|_p = 2.$$

But $2^{1/p} > 2$ if $0 < p < 1$, so the triangle inequality does not hold.

3.2.6

Prove that the space $C[a, b]$ with the metric

$$\rho(f, g) = \int_a^b |f(t) - g(t)| dt$$

is not a complete metric space.

Solution. Note that the integral defining ρ is guaranteed to exist by Theorem 2.11.1 and the extreme value theorem from elementary analysis, and it can be taken in the Riemann sense. It is easily seen that ρ is indeed a metric.

For simplicity, let us assume $a = 0$ and $b = 1$. (The general argument is identical modulo a coordinate transformation.) Define a sequence of functions in $C[0, 1]$ by $f_n(x) = x^n$. Clearly (f_n) converges pointwise to the discontinuous function

$$f(x) = \begin{cases} 0 & (0 \leq x < 1), \\ 1 & (x = 1). \end{cases}$$

For $n \geq m$ we have $f_n \geq f_m$, hence

$$\rho(f_n, f_m) = \int_0^1 f_n(t) dt - \int_0^1 f_m(t) dt = \frac{1}{n+1} - \frac{1}{m+1},$$

which goes to zero as $n, m \rightarrow \infty$. Thus (f_n) is a Cauchy sequence in $C[0, 1]$ which converges to a function f which is not in $C[0, 1]$, so $(C[0, 1], \rho)$ is not a complete metric space.

Section 3.4 – Complete Metric Spaces

Problems

3.4.5

A set Y in a metric space X is said to be *of the first category in X* if it is contained in a countable union of nowhere dense sets of X . If Y is not of the first category in X , then it is said to be *of the second category in X* . The real line with the Euclidean metric is a space of the second category. Prove, however, that, as a subset of the Euclidean plane, the real line is a set of the first category.

Solution. We identify the real line with the subset $L = \{(x, y) \in \mathbb{R}^2 : y = 0\}$, with the induced metric topology. Indeed L is homeomorphic to \mathbb{R} via the map $(x, 0) \mapsto x$, as can easily be shown.

The sets of the form $B((0, 0), r) \cap L = \{(x, 0) : |x| < r\}$ are easily seen to be nowhere dense in \mathbb{R}^2 , and we have

$$L = \bigcup_{n=1}^{\infty} B((0, 0), n) \cap L.$$

Hence L is of the first category in \mathbb{R}^2 .

3.4.7

Let $f(x)$ be a real-valued function on the real line. Prove that there is a nonempty interval (a, b) and a positive number c such that for any $x \in (a, b)$ there is a sequence $\{x_n\}$ such that $x_n \rightarrow x$ and $|f(x_n)| \leq c$.

Solution. Note that

$$\mathbb{R} = |f|^{-1}(\mathbb{R}) = |f|^{-1}\left(\bigcup_{n=1}^{\infty} (-\infty, n]\right) = \bigcup_{n=1}^{\infty} |f|^{-1}((-\infty, n]).$$

Since \mathbb{R} is of the second category (by Theorem 3.4.2), the sets $|f|^{-1}((-\infty, n])$ cannot all be nowhere dense. Hence there is a positive integer c such that the closure of $|f|^{-1}((-\infty, c])$ has nonempty interior. On account of being a nonempty open set, this interior contains an open interval (a, b) .

Let $x \in (a, b)$. Then, since x is contained in the closure of $|f|^{-1}((-\infty, c])$, every neighborhood of x contains a point of $|f|^{-1}((-\infty, c])$. It follows that we can construct a sequence (x_n) in $|f|^{-1}((-\infty, c])$ converging to x , and this sequence will satisfy $|f(x_n)| \leq c$ for all n .

Section 3.5 – Compact Metric Spaces

Problems

3.5.4

A subset F of a compact metric space is compact if and only if it is closed.

Solution. Let X be a compact metric space, and F a subset of X . If F is compact then it is also closed by Corollary 3.5.5.

Conversely, suppose that F is closed. Let \mathcal{C} be any open cover of F . The collection $\mathcal{C} \cup \{X - F\}$ is an open cover of X , hence it has a finite subcover, consisting of some sets $E_1, \dots, E_n \in \mathcal{C}$ and perhaps $\{X - F\}$. The sets E_1, \dots, E_n cover F , hence F is compact.

3.5.5

A subset Y of a metric space is totally bounded if and only if its closure \overline{Y} is totally bounded.

Solution. Clearly Y is totally bounded whenever \overline{Y} is. Conversely, suppose Y is totally bounded, and let $\epsilon > 0$ be given. Then Y admits a finite $\epsilon/2$ -covering,

$$\{B(x_1, \epsilon/2), \dots, B(x_n, \epsilon/2)\}.$$

The union of the closed balls,

$$\bigcup_{i=1}^n \overline{B}(x_i, \epsilon/2),$$

is a closed set containing Y , hence it contains \overline{Y} . It follows that

$$\overline{Y} \subset \bigcup_{i=1}^n \overline{B}(x_i, \epsilon/2) \subset \bigcup_{i=1}^n B(x_i, \epsilon),$$

so \overline{Y} admits a finite ϵ -covering.

3.5.6

The intersection of any number of compact subsets of a metric space is a compact space.

Solution. Let \mathcal{C} be a collection of compact subsets of a metric space X . By Corollary 3.5.5 each $C \in \mathcal{C}$ is closed, hence the intersection $K = \bigcap \mathcal{C}$ is closed as well. Let C' be some particular member of \mathcal{C} . Then K is a closed subset of the compact space C' , thus itself compact by Problem 3.5.4.

Remark: Really the above shows that K is compact *in* C' . However, one easily proves the general result that if F is compact in G , and G is compact in H , then F is compact in H . Hence K is indeed compact in X .

3.5.8

Show that a metric space is compact if and only if it has the following property: for every collection of closed subsets $\{F_\alpha\}$, if any finite subcollection has nonempty intersection, then the whole collection has a nonempty intersection.

Solution. The wording of the problem is potentially misleading. Replace “any” with “every” to make it nonambiguous. Also, let us introduce some useful terminology. A collection of subsets of a topological space is said to have the *finite intersection property* iff every finite subcollection has nonempty intersection. Our task is therefore to show that a metric space X is compact iff every collection \mathcal{F} of closed subsets of X having the finite intersection property has nonempty intersection. However, the proof we will use is valid for general topological spaces, not just metric spaces.

Let X be a topological space, and suppose first that X is compact. Let \mathcal{F} be a collection of closed subsets of X with the finite intersection property. Assume towards a contradiction that $\bigcap \mathcal{F} = \emptyset$. Then

$$\bigcup_{F \in \mathcal{F}} (X - F) = X - \bigcap_{F \in \mathcal{F}} F = X,$$

so $\{X - F; F \in \mathcal{F}\}$ is an open cover of X . Since X is compact, there are sets $F_1, \dots, F_n \in \mathcal{F}$ such that

$$X = \bigcup_{i=1}^n (X - F_i) = X - \bigcap_{i=1}^n F_i.$$

But then $\bigcap_{i=1}^n F_i = \emptyset$, contradicting the hypothesis that \mathcal{F} has the finite intersection property. We conclude that $\bigcap \mathcal{F} \neq \emptyset$.

Conversely, suppose that X has the property described in the problem statement: every collection of closed subsets of X with the finite intersection property has nonempty intersection. Let \mathcal{U} be any open cover of X , and assume towards a contradiction that \mathcal{U} has no finite subcover. Then

$$\bigcap_{i=1}^n (X - U_i) = X - \bigcup_{i=1}^n U_i \neq \emptyset$$

for every finite subcollection $\{U_1, \dots, U_n\} \subset \mathcal{U}$. Hence $\mathcal{F} = \{X - U; U \in \mathcal{U}\}$ is a collection of closed subsets with the finite intersection property. By hypothesis \mathcal{F} has nonempty intersection, so

$$\emptyset = X - \bigcup_{U \in \mathcal{U}} U = \bigcap_{U \in \mathcal{U}} (X - U) = \bigcap_{F \in \mathcal{F}} F \neq \emptyset,$$

a contradiction. We conclude that \mathcal{U} does indeed have a finite subcover, and that X is compact.

Chapter 4 – Elements of Functional Analysis in Banach Spaces

Section 4.1 – Linear Normed Spaces

Problems

4.1.4

If $\{x_n\}$ is a convergent sequence in a normed linear space, with limit x , then also the sequence with elements $(x_1 + \cdots + x_n)/n$ is convergent to x .

Solution. For $n = 1, 2, \dots$, define

$$\sigma_n = \frac{x_1 + \cdots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Let $\epsilon > 0$ be given. Since (x_n) is convergent there exists a positive integer m such that $\|x_n - x\| < \epsilon/2$ for all $n \geq m + 1$. For such n we have

$$\begin{aligned} \|\sigma_n - x\| &= \left\| \frac{1}{n} \sum_{i=1}^n (x_i - x) \right\| \\ &\leq \frac{1}{n} \left\| \sum_{i=1}^m (x_i - x) \right\| + \frac{1}{n} \left\| \sum_{i=m+1}^n (x_i - x) \right\| \\ &\leq \frac{1}{n} \left\| \sum_{i=1}^m (x_i - x) \right\| + \frac{1}{n} \sum_{i=m+1}^n \|x_i - x\| \\ &\leq \frac{1}{n} \left\| \sum_{i=1}^m (x_i - x) \right\| + \frac{n - m - 1}{n} \cdot \frac{\epsilon}{2}. \end{aligned}$$

The first term on the right-hand side goes to zero as $n \rightarrow \infty$, and the other goes to $\epsilon/2$. Hence $\|\sigma_n - x\| < \epsilon$ for sufficiently large n , and it follows that $\sigma_n \rightarrow x$.

4.1.6

A normed linear space is a Banach space if the following property is satisfied: every absolutely convergent series is convergent.

Solution. Let X be a normed linear space with the property that every absolutely convergent series is convergent. Let (x_n) be a Cauchy sequence in X , and choose integers $N_1 < N_2 < \dots$ such that $m, n \geq N_k$ implies $\|x_m - x_n\| < 2^{-k}$. Define a new sequence (y_k) by $y_1 = x_{N_1}$ and $y_k = x_{N_k} - x_{N_{k-1}}$ for $k > 1$. Then

$$\sum_{k=1}^{\infty} \|y_k\| = \|x_{N_1}\| + \sum_{k=2}^{\infty} \|x_{N_k} - x_{N_{k-1}}\| \leq \|x_{N_1}\| + \sum_{k=1}^{\infty} 2^{-k} = \|x_{N_1}\| + 1.$$

It follows by our hypothesis on X that

$$x_{N_k} = \sum_{i=1}^k y_i$$

converges to some $x \in X$. Given $\epsilon > 0$, choose an integer $j \geq 1$ such that $2^{-j} < \epsilon$, and $k > j$ such that

$$\|x_{N_k} - x\| < \epsilon - 2^{-j}.$$

Then, for all $n \geq N_k$,

$$\|x_n - x\| \leq \|x_n - x_{N_k}\| + \|x_{N_k} - x\| < 2^{-k} + \epsilon - 2^{-j} < \epsilon.$$

We conclude that every Cauchy sequence in X is convergent, and hence that X is a Banach space.

Section 4.2 – Subspaces and Bases

Problems

4.2.6

If a linear vector space is infinite-dimensional, then there exist on it norms that are not equivalent. [*Hint:* Let $\{y_\alpha\}$ be a Hamel basis and define norms by $\|x\|^2 = \sum_{\alpha} c_{\alpha} |\lambda_{\alpha}|^2$, where x has the form (4.2.2) and c_{α} are positive numbers.]

Solution. Let $\{y_{\alpha}\}_{\alpha \in I}$ be a Hamel basis for an infinite-dimensional linear space X . For every $x = \sum_{\alpha \in I} \lambda_{\alpha} y_{\alpha}$, define

$$\|x\|_1 = \left(\sum_{\alpha \in I} |\lambda_{\alpha}|^2 \right)^{1/2}.$$

Then $\|\cdot\|_1$ is easily seen to be a norm on X .

Let $\{\alpha_i\}_{i=1}^\infty$ be a countable subset of the index set I . Define a set $\{c_\alpha\}_{\alpha \in I}$ of positive constants by $c_{\alpha_i} = 2^{-i}$ for $i = 1, 2, \dots$, and $c_\alpha = 1$ for $\alpha \notin \{\alpha_i\}$. Define another norm $\|\cdot\|_2$ by

$$\|x\|_2 = \left(\sum_{\alpha \in I} c_\alpha |\lambda_\alpha|^2 \right)^{1/2},$$

for $x = \sum_{\alpha \in I} \lambda_\alpha y_\alpha$.

To see that these two norms are not equivalent, note that

$$\|y_{\alpha_i}\|_2^2 = c_{\alpha_i} = 2^{-i} = 2^{-i} \|y_{\alpha_i}\|_1^2$$

for all i . Hence there exists no $\beta > 0$ such that $\|x\|_1 \leq \beta \|x\|_2$ for all $x \in X$.

Section 4.3 – Finite-Dimensional Normed Linear Spaces

Problems

4.3.1

Let X be a finite-dimensional linear space. Then any two norms on X are equivalent. (According to Problem 4.2.6, the assertion is false if X is infinite-dimensional.)

Solution. Let e_1, \dots, e_n be a basis of X . For every $x = \sum_{i=1}^n \lambda_i e_i \in X$, let

$$\|x\|_1 = \sum_{i=1}^n |\lambda_i|.$$

It is easily verified that $\|\cdot\|_1$ is a norm on X . We will show that every norm on X is equivalent to $\|\cdot\|_1$, and hence (by transitivity) that any two norms on X are equivalent.

Given an arbitrary norm $\|\cdot\|$, we must prove the existence of positive constants α and β such that

$$\alpha \|x\|_1 \leq \|x\| \leq \beta \|x\|_1$$

for all $x \in X$. These inequalities hold trivially for $x = 0$, so it suffices to consider nonzero x . In fact it is sufficient to consider x in the “ $\|\cdot\|_1$ -sphere” $S_1 = \{x \in X : \|x\|_1 = 1\}$, where the inequalities reduce to

$$\alpha \leq \|x\| \leq \beta.$$

The inequality for general, nonzero x then follows upon division by $\|x\|_1$.

We start by showing that the map $x \mapsto \|x\|$ is continuous with respect to the metric $\rho_1(x, y) = \|x - y\|_1$. Let $\epsilon > 0$ be given, and write $M = \max(\|e_1\|, \dots, \|e_n\|)$. Given

$$x = \sum_{i=1}^n \lambda_i e_i, \quad y = \sum_{i=1}^n \mu_i e_i$$

satisfying $\rho_1(x, y) < \epsilon/M$, we have

$$\|x - y\| \leq \sum_{i=1}^n |\lambda_i - \mu_i| \|e_i\| \leq \sum_{i=1}^n |\lambda_i - \mu_i| M = \rho_1(x, y) M < \epsilon,$$

and continuity follows.

The sphere S_1 is obviously closed and bounded under $\|\cdot\|_1$, hence compact by Theorem 4.3.3. By Theorem 3.6.2 and the continuity established above, the map $x \mapsto \|x\|$ attains a maximum and a minimum on S_1 . Let

$$\alpha = \inf_{x \in S_1} \|x\|, \quad \beta = \sup_{x \in S_1} \|x\|,$$

and note that $\alpha > 0$ since $\|x\| = \alpha$ is attained for some nonzero x . It follows that α and β are positive constants such that $\alpha \leq \|x\| \leq \beta$ for all $x \in S_1$, so we are done.

4.3.2

Let Y be a finite-dimensional linear subspace of a normed linear space X , and let $x_0 \in X$, $x_0 \notin Y$. Then there exists a point $y_0 \in Y$ such that

$$\inf_{y \in Y} \|x_0 - y\| = \|x_0 - y_0\|.$$

Solution. Note that Y is closed by Theorem 4.3.2. Let $L = \inf_{y \in Y} \|x_0 - y\|$. For $n = 1, 2, \dots$, choose $y_n \in Y$ such that

$$\|x_0 - y_n\| < L + \frac{1}{n}.$$

Note that $y_n \in B(x_0, L + 1) \cap Y$ for all n . This is a bounded subset of Y , so by Theorems 4.3.3 and 3.5.4, the sequence (y_n) has a subsequence (y_{n_k}) converging to a point

$$y_0 \in \overline{B(x_0, L + 1) \cap Y} \subset \overline{Y} = Y.$$

For all k we have

$$\|x_0 - y_0\| \leq \|x_0 - y_{n_k}\| + \|y_{n_k} - y_0\|.$$

The left-hand side of this inequality is bounded below by L , and the right-hand side converges to L as $k \rightarrow \infty$. It follows that $\|x_0 - y_0\| = L$.

4.3.3

A norm $\| \cdot \|$ is said to be *strictly convex* if $\|x\| = 1$, $\|y\| = 1$, $\|x + y\| = 2$ imply that $x = y$. Prove that if the norm of X is strictly convex, then the point y_0 occurring in the assertion of Problem 4.3.2 is unique.

Solution. Let $\| \cdot \|$ be a strictly convex norm on a linear space X , and let Y be a finite-dimensional linear subspace of X . Let $x_0 \in X - Y$, $L = \inf_{y \in Y} \|x_0 - y\|$, and suppose that there are elements $y_0, y'_0 \in Y$ such that

$$\|x_0 - y_0\| = \|x_0 - y'_0\| = L.$$

It is clear that $\|(x_0 - y_0)/L\| = \|(x_0 - y'_0)/L\| = 1$. Moreover,

$$\|(x_0 - y_0)/L + (x_0 - y'_0)/L\| = \frac{2}{L}\|x_0 - (y_0 + y'_0)/2\| \geq \frac{2}{L}L = 2,$$

and

$$\|(x_0 - y_0)/L + (x_0 - y'_0)/L\| \leq \frac{1}{L}(\|x_0 - y_0\| + \|x_0 - y'_0\|) = \frac{1}{L}2L = 2,$$

so $\|(x_0 - y_0)/L + (x_0 - y'_0)/L\| = 2$. By strict convexity,

$$\frac{1}{L}(x_0 - y_0) = \frac{1}{L}(x_0 - y'_0),$$

and it follows that $y_0 = y'_0$.

4.3.4

Prove that the norm of $L^p(X, \mu)$ is strictly convex if $1 < p < \infty$, and is not strictly convex if $p = 1$ or if $p = \infty$.

Solution. Let $f, g \in \mathcal{L}^p(X, \mu)$, $1 < p < \infty$, and suppose $\|f\|_p = \|g\|_p = 1$ and $\|f + g\|_p = 2$. We then have equality in Minkowski's inequality:

$$\|f + g\|_p = \|f\|_p + \|g\|_p.$$

By Problem 3.2.7 this implies that $f = 0$ a.e., or $g = 0$ a.e., or $f = \lambda g$ a.e. for some positive constant λ . The first two possibilities are ruled out since $\|f\|_p = \|g\|_p = 1$, so the third alternative must hold. But then

$$1 = \|f\|_p = |\lambda| \|g\|_p = |\lambda|,$$

so $\lambda = 1$. Thus $f = g$ a.e., so that $\tilde{f} = \tilde{g}$ in $L^p(X, \mu)$. It follows that $\| \cdot \|_p$ is strictly convex if $1 < p < \infty$.

For $p = 1, \infty$, surely the problem is supposed to say 'not *necessarily* strictly convex', because we can come up with examples where $\| \cdot \|_p$ is strictly convex, such as an empty measure space $X = \emptyset$, or any space with identically zero measure $\mu = 0$. (Perhaps less trivial examples exist.) Hence we will only

demonstrate that there are examples where $\|\cdot\|_p$ ($p \in \{1, \infty\}$) is not strictly convex.

For $p = 1$, consider the L^1 -space l^1 . The sequences $x = (1, 0, 0, \dots)$ and $y = (0, 1, 0, 0, \dots)$ satisfy $\|x\|_1 = \|y\|_1 = 1$ and $\|x + y\|_1 = 2$, yet $x \neq y$. For $p = \infty$, consider the L^∞ -space l^∞ . The sequences $x = (1, 0, 0, \dots)$ and $y = (1, 1, 1, \dots)$ satisfy $\|x\|_\infty = \|y\|_\infty = 1$ and $\|x + y\|_\infty = 2$, but $x \neq y$.

4.3.5

Prove that in $C[a, b]$ the uniform norm is not equivalent to the L^p norm (for $1 \leq p < \infty$).

Solution. For simplicity, let $a = 0$ and $b = 1$. For $n = 1, 2, \dots$, define $f_n : [0, 1] \rightarrow \mathbb{R}$ (or \mathbb{C}) by $f_n(x) = x^n$. Let $\|\cdot\|_u$ denote the uniform norm. For all n we have

$$\|f_n\|_u = \max_{0 \leq x \leq 1} |f_n(x)| = 1$$

and

$$\|f_n\|_p = \left(\int_0^1 x^n dx \right)^{1/p} = (n+1)^{-1/p}.$$

(We are assuming that the L^p norm is defined with the standard Lebesgue measure.) In particular $\|f_n\|_p \rightarrow 0$ as $n \rightarrow \infty$, whatever the value of p ($\neq \infty$). It follows that there exists no $\beta > 0$ such that $\|f\|_u \leq \beta \|f\|_p$ for all $f \in C[0, 1]$, and hence that the two norms are not equivalent.

4.3.7

Let n be a positive integer, $1 \leq p < \infty$, and let $f(x)$ be a continuous function on $0 \leq x \leq 1$. Then there exists a unique polynomial Q_n of degree n such that for any other polynomial P_n of degree n

$$\int_0^1 |f(x) - P_n(x)|^p dx > \int_0^1 |f(x) - Q_n(x)|^p dx.$$

Solution. There is an error in the problem statement; it should be polynomials of degree $\leq n$, not *exactly* n . To see that the written claim is false, consider the case $f(x) = 0$ and $n = 1$. Then f can be approximated arbitrarily closely by degree 1 polynomials $ax + b$, $a \neq 0$, but no such polynomial will make $\|f - P\|_p$ vanish completely.

Let \mathcal{P}_n denote the set of polynomial functions on $[0, 1]$ with degree $\leq n$. Then \mathcal{P}_n is a linear subspace of the normed linear space $(C[0, 1], \|\cdot\|_p)$, and is finite-dimensional since it is spanned by the polynomials $1, x, x^2, \dots, x^n$.

If $f \in \mathcal{P}_n$, then $Q = f$ satisfies the claim, since $\|f - P\|_p > 0$ for all $P \neq f$ (else $\|\cdot\|_p$ would not be a norm). If $f \notin \mathcal{P}_n$, then the conclusion of Problem 4.3.2 tells us that there exists $Q \in \mathcal{P}_n$ such that

$$\inf_{P \in \mathcal{P}_n} \|f - P\| = \|f - Q\|.$$

This Q is unique by Problem 4.3.3 if the norm $\| \cdot \|_p$ is strictly convex, whereupon the claim follows. By Problem 4.3.4 this is the case for $1 < p < \infty$. In fact the norm is strictly convex even for $p = 1$ over $C[0, 1]$, as is easily verified directly from the definition.

Section 4.4 – Linear Transformations

Problems

4.4.2

Let T be an additive operator [that is, $T(x_1 + x_2) = Tx_1 + Tx_2$] from a real normed linear space X into a normed linear space Y . If T is continuous, then T is homogeneous [that is, $T(\lambda x) = \lambda Tx$]. [*Hint*: Prove that $T[(m/n)x] = (m/n)Tx$, where m, n are integers.]

Solution. Let x be an arbitrary element of X . By induction, additivity implies $T(mx) = mTx$ for positive integers m . Moreover,

$$T0 = T(0 + 0) = T0 + T0$$

implies that $T0 = 0$, so that

$$0 = T0 = T(mx - mx) = T(mx) + T(-mx),$$

which shows that $T(-mx) = -T(mx) = -mTx$, thus extending the earlier result to nonpositive integers. Finally, if m and n are integers, $n \neq 0$, then

$$nT\left(\frac{m}{n}x\right) = T\left(n\frac{m}{n}x\right) = T(mx) = mTx,$$

so that $T[(m/n)x] = (m/n)Tx$, further extending the result to rationals.

Now, let $\lambda \in \mathbb{R}$. There exists a sequence (λ_n) in \mathbb{Q} such that $\lambda_n \rightarrow \lambda$. Clearly $\lambda_n x \rightarrow \lambda x$ in X , so $T(\lambda_n x) \rightarrow T(\lambda x)$ in Y by continuity of T . But, by what we found above,

$$T(\lambda_n x) = \lambda_n Tx \rightarrow \lambda Tx,$$

so $T(\lambda x) = \lambda Tx$.

4.4.6

Find the norm of the operator $A \in \mathcal{B}(X)$ given by $(Af)(t) = tf(t)$ ($0 \leq t \leq 1$), where (a) $X = C[0, 1]$, (b) $X = L^p(0, 1)$ and $(1 \leq p \leq \infty)$.

Solution.

(a) For $0 \leq t \leq 1$ we have $|tf(t)| = |t||f(t)| \leq |f(t)|$, hence $\|Af\| \leq \|f\|$, and

$$\|A\| = \sup_{f \neq 0} \frac{\|Af\|}{\|f\|} \leq 1.$$

The upper bound $\|Af\|/\|f\| = 1$ is attained with f constant, so $\|A\| = 1$.

- (b) Let us first verify that $A \in \mathcal{B}(X)$, i.e. that it is a bounded linear operator $L^p(0, 1) \rightarrow L^p(0, 1)$. Linearity is immediate. If $f \in \mathcal{L}^p(0, 1)$, then $|f|^p$ is integrable, so $|tf(t)|^p = |t|^p|f(t)|^p$ is integrable by Corollary 2.10.2, and we see that A does indeed map into $L^p(0, 1)$. Finally, since $|tf(t)|^p = |t|^p|f(t)|^p \leq |f(t)|^p$ for $0 < t < 1$, we have $\|Af\|_p \leq \|f\|_p$, so that $\|A\| \leq 1$.

We will now show that $\|A\| \geq 1$, so that $\|A\| = 1$. The case $p = \infty$ is similar to (a), so we will assume $1 \leq p < \infty$. For $n = 1, 2, \dots$, define simple functions $f_n : (0, 1) \rightarrow \mathbb{R}$ (or \mathbb{C}) by

$$f_n(t) = \begin{cases} 0 & \text{if } 0 < t < 1 - 1/n, \\ n^{1/p} & \text{if } 1 - 1/n \leq t < 1. \end{cases}$$

Then one easily finds that $\|f_n\|_p = 1$ and $\|Af_n\|_p \geq 1 - 1/n$, so that

$$\|A\| \geq \frac{\|Af_n\|_p}{\|f_n\|_p} \geq 1 - \frac{1}{n}$$

for all n . Indeed it follows that $\|A\| \geq 1$.

4.4.7

A linear operator from a normed linear space X into a normed linear space Y is bounded if and only if it maps bounded sets onto bounded sets.

Solution. Let $T : X \rightarrow Y$ be a linear operator between normed linear spaces.

Suppose first that T is bounded, and let A be a bounded subset of X . Then $\|T\| < \infty$, and there is some L such that $\|x\| \leq L < \infty$ for all $x \in A$. Hence

$$\|Tx\| \leq \|T\| \|x\| \leq \|T\| L$$

for all $x \in A$, so $T(A)$ is bounded.

Conversely, suppose that T maps bounded sets onto bounded sets. The set $\{x \in X : \|x\| = 1\}$ is bounded, so there exists M such that $\|Tx\| \leq M < \infty$ whenever $\|x\| = 1$. It follows that

$$\|T\| = \sup_{\|x\|=1} \|Tx\| \leq M,$$

and hence that T is bounded.

4.4.8

A linear operator from a normed linear space X into a normed linear space Y is continuous if and only if it maps sequences converging to 0 into bounded sequences.

Solution. Let $T : X \rightarrow Y$ be a linear operator between normed linear spaces. The claim is immediate if X is the trivial space (containing only the 0 vector), so let us assume that X is nontrivial.

Suppose first that T is continuous, and therefore bounded. Any sequence in X converging to 0 is easily seen to be bounded, hence is mapped to a bounded sequence by the conclusion of the previous problem.

Conversely, suppose T has the property that it maps sequences converging to 0 into bounded sequences. Assume towards a contradiction that T is unbounded. Then it is possible to construct a sequence (x_n) in X such that $\|x_n\| = 1$ and $\|Tx_n\| > n$ for every n . The sequence (x_n/\sqrt{n}) converges to 0, so $\{T(x_n/\sqrt{n})\}$ is bounded by hypothesis. But

$$\left\| T \left(\frac{x_n}{\sqrt{n}} \right) \right\| = \frac{1}{\sqrt{n}} \|Tx_n\| > \frac{1}{\sqrt{n}} \cdot n = \sqrt{n} \rightarrow \infty$$

as $n \rightarrow \infty$, yielding a contradiction. We conclude that T must be bounded, and thus continuous.

Section 4.6 – The Open-Mapping Theorem and the Closed-Graph Theorem

Problems

4.6.1

If T, S, T^{-1}, S^{-1} belong to $\mathcal{B}(X)$, then $(TS)^{-1} \in \mathcal{B}(X)$ and $(TS)^{-1} = S^{-1}T^{-1}$.

Solution. Note in particular that T, S, T^{-1}, S^{-1} belonging to $\mathcal{B}(X)$ implies that all of these transformations are bijective, in addition to bounded. (If T is not surjective, then $D_{T^{-1}} = T(X) \neq X$, hence $T^{-1} \notin \mathcal{B}(X)$.)

It is clear that $\|TS\| \leq \|T\| \|S\| < \infty$, hence $TS \in \mathcal{B}(X)$, and it is bijective on account of being a composition of bijections. Thus $(TS)^{-1}$ exists and is equal to $S^{-1}T^{-1}$. It follows that $\|(TS)^{-1}\| \leq \|S^{-1}\| \|T^{-1}\| < \infty$, hence $(TS)^{-1} \in \mathcal{B}(X)$.

4.6.2

Let X be a Banach space and let $A \in \mathcal{B}(X)$, $\|A\| < 1$. Prove that $(I + A)^{-1}$ exists and is given by

$$(I + A)^{-1} = \sum_{n=0}^{\infty} (-1)^n A^n,$$

where the series is absolutely convergent [in $\mathcal{B}(X)$]. Show also that

$$\|(I + A)^{-1}\| \leq 1/(1 - \|A\|).$$

Solution. We note first that

$$\sum_{n=0}^{\infty} \|(-1)^n A^n\| = \sum_{n=0}^{\infty} \|A^n\| \leq \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|},$$

since $\|A\| < 1$. Hence the series $\sum_{n=0}^{\infty} (-1)^n A^n$ is strongly convergent to an operator $B \in \mathcal{B}(X)$, by Theorem 4.5.2. (Recall that $\mathcal{B}(X)$ is a Banach space, hence absolute convergence implies convergence. Convergence in $\mathcal{B}(X)$ is the uniform convergence of operators, and uniform convergence implies strong convergence.)

For all $x \in X$, we have

$$ABx = A \left(\sum_{n=0}^{\infty} (-1)^n A^n x \right) = \sum_{n=0}^{\infty} (-1)^n A^{n+1} x = x - \sum_{n=0}^{\infty} (-1)^n A^n x = (I - B)x,$$

where we have used the continuity of A to interchange limiting processes. Also,

$$BAx = \sum_{n=0}^{\infty} (-1)^n A^{n+1} x = ABx.$$

It follows that

$$B(I + A) = (I + A)B = B + AB = B + I - B = I,$$

so that $B = (I + A)^{-1}$.

Finally,

$$\|(I + A)^{-1}\| \leq \sum_{n=0}^{\infty} \|(-1)^n A^n\| \leq \frac{1}{1 - \|A\|}$$

by what we found earlier.

4.6.3

Let X be a Banach space and let T and T^{-1} belong to $\mathcal{B}(X)$. Prove that if $S \in \mathcal{B}(X)$ and $\|S - T\| < 1/\|T^{-1}\|$, then S^{-1} exists and is a bounded operator, and

$$\|S^{-1} - T^{-1}\| < \frac{\|T^{-1}\|}{1 - \|S - T\| \|T^{-1}\|}.$$

[Hint: $S = [(S - T)T^{-1} + I]T$.]

Solution. Note that $ST^{-1} = I + (S - T)T^{-1}$, and that

$$\|(S - T)T^{-1}\| \leq \|S - T\| \|T^{-1}\| < 1.$$

By the previous problem $(ST^{-1})^{-1}$ exists, and

$$\|(ST^{-1})^{-1}\| \leq \frac{1}{1 - \|(S - T)T^{-1}\|} \leq \frac{1}{1 - \|S - T\| \|T^{-1}\|}.$$

Moreover, since $ST^{-1}(ST^{-1})^{-1} = I$ and

$$T^{-1}(ST^{-1})^{-1}S = T^{-1}(ST^{-1})^{-1}ST^{-1}T = T^{-1}T = I,$$

we have $S^{-1} = T^{-1}(ST^{-1})^{-1}$. Finally,

$$\|S^{-1}\| = \|T^{-1}(ST^{-1})^{-1}\| \leq \|T^{-1}\| \|(ST^{-1})^{-1}\| \leq \frac{\|T^{-1}\|}{1 - \|(S - T)\| \|T^{-1}\|},$$

and

$$\|S^{-1} - T^{-1}\| = \|S^{-1}(S - T)T^{-1}\| \leq \|S^{-1}\| \|S - T\| \|T^{-1}\| < \|S^{-1}\|,$$

so that

$$\|S^{-1} - T^{-1}\| < \frac{\|T^{-1}\|}{1 - \|(S - T)\| \|T^{-1}\|}.$$

4.6.4

Let X and Y be two linear vector spaces. Find necessary and sufficient conditions for a subset G of $X \times Y$ to be the graph of a linear operator from X into Y .

Solution. We claim that G is the graph of a linear operator from X into Y if and only if

- (i) G is a linear subspace of $X \times Y$.
- (ii) The set $G \cap (\{0\} \times Y)$ is a singleton.

It is clear that these conditions are necessary, so we need only prove that they are sufficient.

Since G is nonempty by (ii), it contains an element (x, y) . By (i) it also contains $(0, 0) = 0 \cdot (x, y)$, and it follows that $G \cap (\{0\} \times Y) = \{(0, 0)\}$.

If $(u, v), (u, v') \in G$, then

$$(0, v - v') = (u, v) - (u, v') \in G \cap (\{0\} \times Y)$$

by (i). It follows that $(0, v - v') = (0, 0)$, hence that $v = v'$.

By the above, G is a functional relation on $X \times Y$, so it defines a partial function $T : D \subset X \rightarrow Y$, where

$$D = \{x \in X; \exists y \in Y \text{ such that } (x, y) \in G\},$$

and $T(x) = y$ if $(x, y) \in G$.

Suppose $x_1, x_2 \in D$, and λ_1, λ_2 are scalars. Then

$$(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 T(x_1) + \lambda_2 T(x_2)) = \lambda_1 (x_1, T(x_1)) + \lambda_2 (x_2, T(x_2)) \in G$$

by (i). Hence $\lambda_1 x_1 + \lambda_2 x_2 \in D$, which shows that D is a linear subspace of $X \times Y$, and

$$T(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 T(x_1) + \lambda_2 T(x_2),$$

so that T is a linear operator.

Finally, note that G is precisely the graph of T . This completes the proof.

4.6.5

Let X and Y be Banach spaces and let T be a bounded linear map from X into Y . If $T(X)$ is of the second category (in Y), then $T(X) = Y$.

Solution. We will use the following lemma:

If W is a linear subspace of a normed space V , and if W contains a nonempty open subset of V , then $W = V$.

To see that this is true, note that W contains an open ball $B(x_0, r)$. Given any $y \in V$, let

$$x = \frac{r}{2\|y\|}y + x_0.$$

Then $x \in B(x_0, r) \subset W$, so that

$$y = \frac{2\|y\|}{r}(x - x_0) \in W$$

by closure under linear combinations.

Now, assume that $T(X)$ is of the second category, and follow the steps of parts (a) and (b) of the proof of Theorem 4.6.1, but with $T(X)$ in place of Y . We find that $T(X)$ contains an open ball (see equation (4.6.2)) and hence that $T(X) = Y$ by our lemma.

4.6.6

Let X and Y be Banach spaces and let T be a linear map from a linear subspace D_T of X into Y . If D_T (in X) and the graph of T (in $X \times Y$) are closed, then T is bounded—that is, $\|Tx\| \leq K\|x\|$ for all $x \in D_T$ (K constant).

Solution. Note first that D_T is complete; this is true for any closed subset of any complete space. Indeed, if (x_n) is a Cauchy sequence in D_T , then it is also a Cauchy sequence in X , so it converges to a point $x \in X$, and since D_T is closed we have $x \in D_T$. It follows that D_T is a Banach space. Hence we can regard T as a map $D_T \rightarrow Y$ and apply Theorem 4.6.4 to conclude that T is continuous, thus bounded by Theorem 4.4.2.

4.6.7

Let X be a normed linear space with any one of two norms $\|\cdot\|_1, \|\cdot\|_2$. If $\|x_n\|_2 \rightarrow 0$ implies $\|x_n\|_1 \rightarrow 0$, then (4.6.5) holds.

Solution. Assume towards a contradiction that (4.6.5) does not hold. Then, for every $n \in \{1, 2, \dots\}$, there exists $x_n \in X$ such that $\|x_n\|_1 > n\|x_n\|_2$. The sequence $y_n = x_n/(n\|x_n\|_2)$ satisfies

$$\|y_n\|_1 = \frac{\|x_n\|_1}{n\|x_n\|_2} > \frac{n\|x_n\|_2}{n\|x_n\|_2} = 1$$

for all n . But $\|y_n\|_2 = 1/n \rightarrow 0$, which by hypothesis implies $\|y_n\|_1 \rightarrow 0$, yielding a contradiction.

Section 4.8 – The Hahn-Banach Theorem

Problems

4.8.1

Let X be a normed linear space and let $\{x_n\} \subset X$. A point y_0 is the limit of linear combinations $\sum_{j=1}^n c_j x_j$ if and only if $x^*(y_0) = 0$ for all x^* for which $x^*(x_j) = 0$ for $1 \leq j < \infty$.

Solution. Note that “ y_0 is the limit of linear combinations $\sum_{j=1}^n c_j x_j$ ” is not meant to imply $y_0 = \lim_{n \rightarrow \infty} \sum_{j=1}^n c_j x_j$ for some sequence of coefficients (c_j) . Rather, it just says that y_0 is the limit of a sequence of finite linear combinations of elements in $\{x_n\}$; i.e., that y_0 is a limit point of $S := \text{span}\{x_n\}$. Also, the statement “ $x^*(x_j) = 0$ for $1 \leq j < \infty$ ” can be simplified to “ x^* vanishes on S .”

Suppose that y_0 is a limit point of S . Then there is a sequence (s_n) in S such that $s_n \rightarrow y_0$. If $x^* \in X^*$ vanishes on S , then

$$x^*(y_0) = x^*\left(\lim_{n \rightarrow \infty} s_n\right) = \lim_{n \rightarrow \infty} x^*(s_n) = 0$$

by continuity of x^* .

Conversely, suppose that $x^*(y_0) = 0$ for all $x^* \in X^*$ that vanish on S . Let

$$d = \inf_{s \in S} \|s - y_0\|.$$

If $d > 0$, then Theorem 4.8.3 tells us that there exists $x^* \in X^*$ such that $x^*(y_0) = 1$, but which vanishes on S . This contradicts our assumptions, so we conclude that $d = 0$, which in turn shows that y_0 is a limit point of S .

4.8.5

Let X be an infinite-dimensional Banach space. Prove that there exists an infinite, strictly decreasing sequence $\{Y_n\}$ of infinite-dimensional closed linear subspaces of X . [*Hint:* Take Y_1 to be the null space of some $x_1^* \neq 0$ in X^* . Take Y_2 to be the null space of some $x_2^* \neq 0$ in Y_1^* , and so on.]

Solution. We will construct such a sequence by induction. For the base case, let $Y_0 = X$. Certainly Y_0 is an infinite-dimensional closed linear subspace of X .

For the inductive step, assume that we have infinite-dimensional closed linear subspaces $Y_0 \supset Y_1 \supset \cdots \supset Y_n$, with the inclusions being strict. Fix an element $x_0 \in Y_n$ such that $\|x_0\| = 1$. (Such an element is guaranteed to exist since Y_n is infinite-dimensional.) Corollary 4.8.4 provides a continuous linear functional $x^* \in Y_n^*$ such that $x^*(x_0) = 1$. Let Y_{n+1} be the null space of x^* ; clearly a proper linear subspace of Y_n , hence also of X . As discussed after Corollary 4.8.7, each element $x \in Y_n$ can be written as $x = z + \lambda x_0$, where $\lambda = x^*(x)$ and $z = x - \lambda x_0 \in Y_{n+1}$. It follows that Y_{n+1} is infinite-dimensional. Finally, if (y_i) is a sequence in Y_{n+1} and $y_i \rightarrow y \in X$, then $x^*(y) = \lim_{i \rightarrow \infty} x^*(y_i) = 0$ by continuity, so $y \in Y_{n+1}$, making Y_{n+1} a closed subset of X . This concludes the inductive step.

4.8.9

Let $u(t)$ be a function defined on $a < t < b$ with values in a Banach space X . We say that $u(t)$ is *strongly differentiable at t [on (a, b)]* if $\lim_{h \rightarrow 0} \{[u(t+h) - u(t)]/h\}$ exists [for all $t \in (a, b)$]. The limit is denoted by $du(t)/dt$ and is called the derivative of $u(t)$. For functions $A(t)$ with values in $\mathcal{B}(X)$, if $\lim_{h \rightarrow 0} \{[A(t+h)x - A(t)x]/h\}$ exists for any $x \in X$, then we say that $A(t)$ has a *strong derivative*. If $\lim_{h \rightarrow 0} \{[A(t+h) - A(t)]/h\}$ exists (in the uniform topology), then we say that $A(t)$ is *uniformly differentiable*. Prove that e^{tA} [$A \in \mathcal{B}(X)$] is uniformly differentiable and $de^{tA}/dt = Ae^{tA}$.

Solution. First note that, given any $t \in \mathbb{R}$,

$$\sum_{n=0}^{\infty} \left\| \frac{(tA)^n}{n!} \right\| \leq \sum_{n=0}^{\infty} \frac{(|t|^n \|A\|^n)}{n!} = e^{|t| \|A\|} < \infty,$$

so that the series $\sum_{n=0}^{\infty} (tA)^n/(n!)$ is strongly convergent (to an operator) in $\mathcal{B}(X)$ (c.f. Theorems 4.1.2, 4.5.2). Hence we can define a map $E_A : \mathbb{R} \rightarrow \mathcal{B}(X)$ by $E_A(t) = \sum_{n=0}^{\infty} (tA)^n/(n!)$. Instead of $E_A(t)$ we typically write e^{tA} .

With a little work we find that

$$\frac{e^{(t+h)A} - e^{tA}}{h} - Ae^{tA} = \sum_{n=2}^{\infty} \frac{1}{n!} \frac{(t+h)^n - t^n - nht^{n-1}}{h} A^n.$$

Further simplification with the binomial formula yields

$$\frac{(t+h)^n - t^n - nht^{n-1}}{h} = \sum_{k=2}^n \binom{n}{k} t^{n-k} h^{k-1}.$$

Hence

$$\left\| \frac{e^{(t+h)A} - e^{tA}}{h} - Ae^{tA} \right\| \leq \sum_{n=2}^{\infty} \frac{\|A\|^n}{n!} \sum_{k=2}^n \binom{n}{k} |t|^{n-k} |h|^{k-1}.$$

If $t = 0$, then the right-hand side becomes

$$|h| \cdot \|A\|^2 \sum_{n=0}^{\infty} \frac{(|h| \|A\|)^n}{(n+2)!} \leq |h| \|A\|^2 e^{|h| \|A\|},$$

which goes to zero as $h \rightarrow 0$. If $t \neq 0$, then

$$\sum_{k=2}^n \binom{n}{k} |t|^{n-k} |h|^{k-1} \leq |h| |t|^{n-2} \sum_{k=2}^n \binom{n}{k} = |h| |t|^{n-2} 2^n$$

whenever $|h| \leq |t|$, yielding

$$\left\| \frac{e^{(t+h)A} - e^{tA}}{h} - Ae^{tA} \right\| \leq \frac{|h|}{|t|^2} \sum_{n=2}^{\infty} \frac{(2|t| \|A\|)^n}{n!} \leq \frac{|h|}{|t|^2} e^{2|t| \|A\|}$$

for sufficiently small h , which also goes to zero as $h \rightarrow 0$. It follows that e^{tA} is uniformly differentiable on \mathbb{R} , with derivative Ae^{tA} .

4.8.10

Let X be a real normed linear space, and let $u(t)$ be continuous and strongly differentiable in (a, b) . Then for any $a < \alpha < \beta < b$,

$$\|u(\beta) - u(\alpha)\| \leq (\beta - \alpha) \sup_{\alpha \leq t \leq \beta} \left\| \frac{du(t)}{dt} \right\|.$$

[Hint: Apply x^* to $u(\beta) - u(\alpha)$.]

Solution. If $u(\alpha) = u(\beta)$ then there is nothing to prove, so assume $u(\beta) \neq u(\alpha)$. By Corollary 4.8.4 there is a bounded linear operator $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(u(\beta) - u(\alpha)) = \|u(\beta) - u(\alpha)\|$. Define $f = x^* \circ u : (a, b) \rightarrow \mathbb{R}$. Since x^* is linear and continuous,

$$\lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h} = \lim_{h \rightarrow 0} x^* \left[\frac{u(t+h) - u(t)}{h} \right] = x^* \left[\lim_{h \rightarrow 0} \frac{u(t+h) - u(t)}{h} \right],$$

so that f is differentiable with derivative

$$\frac{df(t)}{dt} = x^* \left[\frac{du(t)}{dt} \right].$$

By the mean value theorem from elementary real analysis, we have

$$f(\beta) - f(\alpha) = (\beta - \alpha) \frac{df(\gamma)}{dt}$$

for some $\gamma \in (\alpha, \beta)$. Thus,

$$\begin{aligned} \|u(\beta) - u(\alpha)\| &= x^*(u(\beta) - u(\alpha)) \\ &= f(\beta) - f(\alpha) \\ &= (\beta - \alpha) \frac{df(\gamma)}{dt} \\ &= (\beta - \alpha) x^* \left[\frac{du(\gamma)}{dt} \right] \\ &\leq (\beta - \alpha) \|x^*\| \cdot \left\| \frac{du(\gamma)}{dt} \right\| \\ &\leq (\beta - \alpha) \sup_{\alpha \leq t \leq \beta} \left\| \frac{du(t)}{dt} \right\|. \end{aligned}$$

4.8.12

For every normed linear space X there is a set A such that X is isomorphic to a subspace of the Banach space of functions f on A with norm $\|f\| = \sup_{t \in A} |f(t)|$. If X is separable, A is countable. [Hint: Let $\{x_\alpha : \alpha \in A\}$ be dense in X . Let $f(x, \alpha)$ be the bounded linear functional (in $x \in X$) satisfying $\|f(\cdot, \alpha)\| = 1$, $f(x_\alpha, \alpha) = \|x_\alpha\|$. Define the isomorphism $x \rightarrow g_x(\alpha)$, where $g_x(\alpha) = f(x, \alpha)$. Prove: $\|f(x, \alpha) - \|x\|\| \leq 2\|x_\alpha - x\|$.]

Solution. The conclusion is immediate if X is a trivial space, so assume $X \neq \{0\}$. Let $\{x_\alpha\}_{\alpha \in A}$ be any dense subset of X (possibly X itself), indexed by a set A . (Recall that any space can be indexed by itself.) For each $\alpha \in A$, Corollary 4.8.4 fashions a bounded linear functional $f_\alpha \in X^*$ such that $\|f_\alpha\| = 1$ and $f_\alpha(x_\alpha) = \|x_\alpha\|$.

For each $x \in X$, define $g_x : A \rightarrow \mathbb{F}$ (where \mathbb{F} is the field associated with X) by $g_x(\alpha) = f_\alpha(x)$. We will prove the following properties of the functions g_x :

- (i) If $x, y \in X$ and $\lambda, \mu \in \mathbb{F}$, then $\lambda g_x + \mu g_y = g_{\lambda x + \mu y}$.
- (ii) Each function g_x is bounded, with $\sup_{\alpha \in A} |g_x(\alpha)| = \|x\|$.

To prove (i), simply let $\alpha \in A$ and compute

$$(\lambda g_x + \mu g_y)(\alpha) = \lambda g_x(\alpha) + \mu g_y(\alpha) = \lambda f_\alpha(x) + \mu f_\alpha(y) = f_\alpha(\lambda x + \mu y) = g_{\lambda x + \mu y}(\alpha),$$

using the linearity of f_α .

Proving (ii) takes more work. Note first that

$$\sup_{\alpha \in A} |g_x(\alpha)| = \sup_{\alpha \in A} |f_\alpha(x)| \leq \sup_{\alpha \in A} \|f_\alpha\| \|x\| = \|x\|.$$

Next, given any $\epsilon > 0$, fix $\alpha' \in A$ such that $\|x_{\alpha'} - x\| < \epsilon/2$. (This is possible because $\{x_\alpha\}_{\alpha \in A}$ is dense in X .) By the triangle inequality,

$$\|x\| \leq \|x\| - \|x_{\alpha'}\| + \|x_{\alpha'}\| - |f_{\alpha'}(x)| + |f_{\alpha'}(x)|.$$

Now, the “reverse triangle inequality” yields

$$\|x\| - \|x_{\alpha'}\| \leq \|x_{\alpha'} - x\| < \frac{\epsilon}{2}$$

and

$$\begin{aligned} |\|x_{\alpha'}\| - |f_{\alpha'}(x)|| &\leq |\|x_{\alpha'}\| - f_{\alpha'}(x)| \\ &= |f_{\alpha'}(x_{\alpha'}) - f_{\alpha'}(x)| \\ &= |f_{\alpha'}(x_{\alpha'} - x)| \\ &\leq \|f_{\alpha'}\| \|x_{\alpha'} - x\| \\ &= \|x_{\alpha'} - x\| \\ &< \frac{\epsilon}{2}. \end{aligned}$$

Also,

$$|f_{\alpha'}(x)| \leq \sup_{\alpha \in A} |f_\alpha(x)| = \sup_{\alpha \in A} |g_x(\alpha)|.$$

Putting it all together, we have

$$\|x\| \leq \sup_{\alpha \in A} |g_x(\alpha)| + \epsilon,$$

and since this holds for every ϵ we finally arrive at (ii).

Let F_A be the Banach space of all bounded functions $A \rightarrow \mathbb{F}$, with the supremum norm $\|g\| = \sup_{\alpha \in A} |g(\alpha)|$. Since the functions g_x are bounded by (ii), we can define a map $\sigma : X \rightarrow F_A$ by $\sigma(x) = g_x$. The image of σ is a linear subspace of F_A by (i). Moreover, σ is an imbedding, since

$$\|\sigma(x) - \sigma(y)\| = \|g_x - g_y\| = \|g_{x-y}\| = \sup_{\alpha \in A} |g_{x-y}(\alpha)| = \|x - y\|$$

for all $x, y \in X$, by (i) and (ii). Every imbedding is injective, so σ is an isomorphism (in the sense of Section 3.3) onto its image.

Finally, if X is separable, then we can take A to be countable.